wwPDB EM Validation Summary Report (i)

Nov 21, 2022 – 06:34 PM JST

PDB ID	:	7D9U
EMDB ID	:	EMD-30621
Title	:	Structure of human soluble guanylate cyclase in the cinciguat-bound activated state
Authors	:	Chen, L.; Liu, R.; Kang, Y.
Deposited on	:	2020-10-14
Resolution	:	3.80 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 (1)) were used in the production of this report:

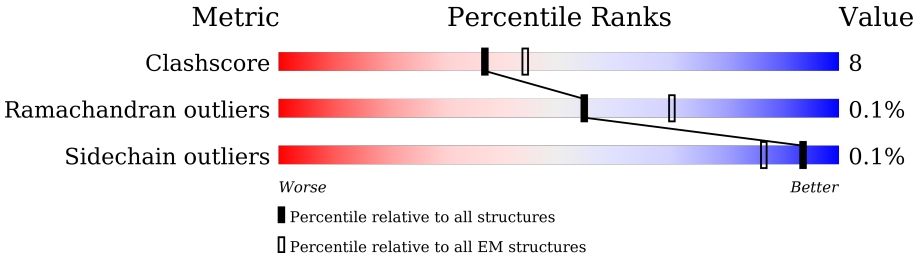
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 3.80 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	690	21% 61% 12% 27%
2	B	619	18% 73% 19% 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8190 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 Guanylate cyclase soluble subunit alpha-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	506	3742	2416	630	667	29	0	0

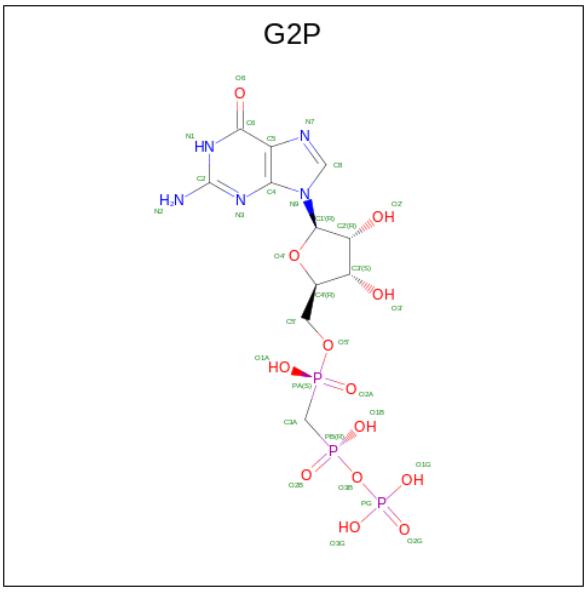
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	44	MET	VAL	variant	UNP Q02108
A	554	VAL	LEU	variant	UNP Q02108

- Molecule 2 is a protein called Guanylate cyclase soluble subunit beta-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	569	4372	2801	753	791	27	2	0

- Molecule 3 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-letter code: G2P) (formula: $C_{11}H_{18}N_5O_{13}P_3$).

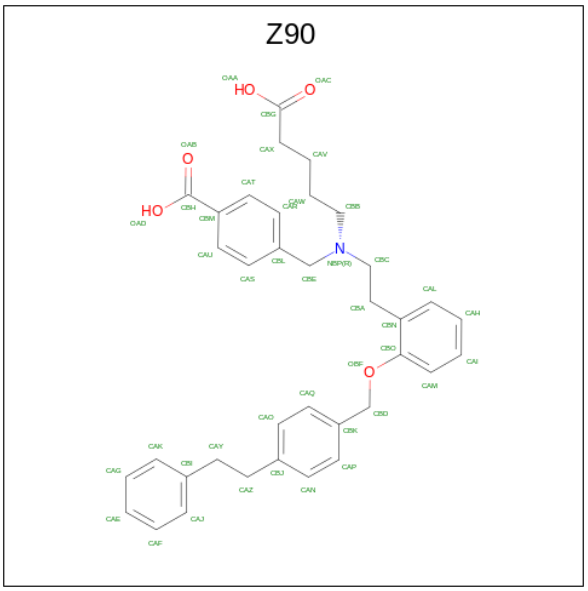


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
3	A	1	32	11	5	13	3	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
4	A	2	2	2	0

- Molecule 5 is 4-({(4-carboxybutyl)[2-(2-{[4-(2-phenylethyl)benzyl]oxy}phenyl)ethyl]amino }methyl)benzoic acid (three-letter code: Z90) (formula: $C_{36}H_{39}NO_5$) (labeled as "Ligand of Interest" by depositor).



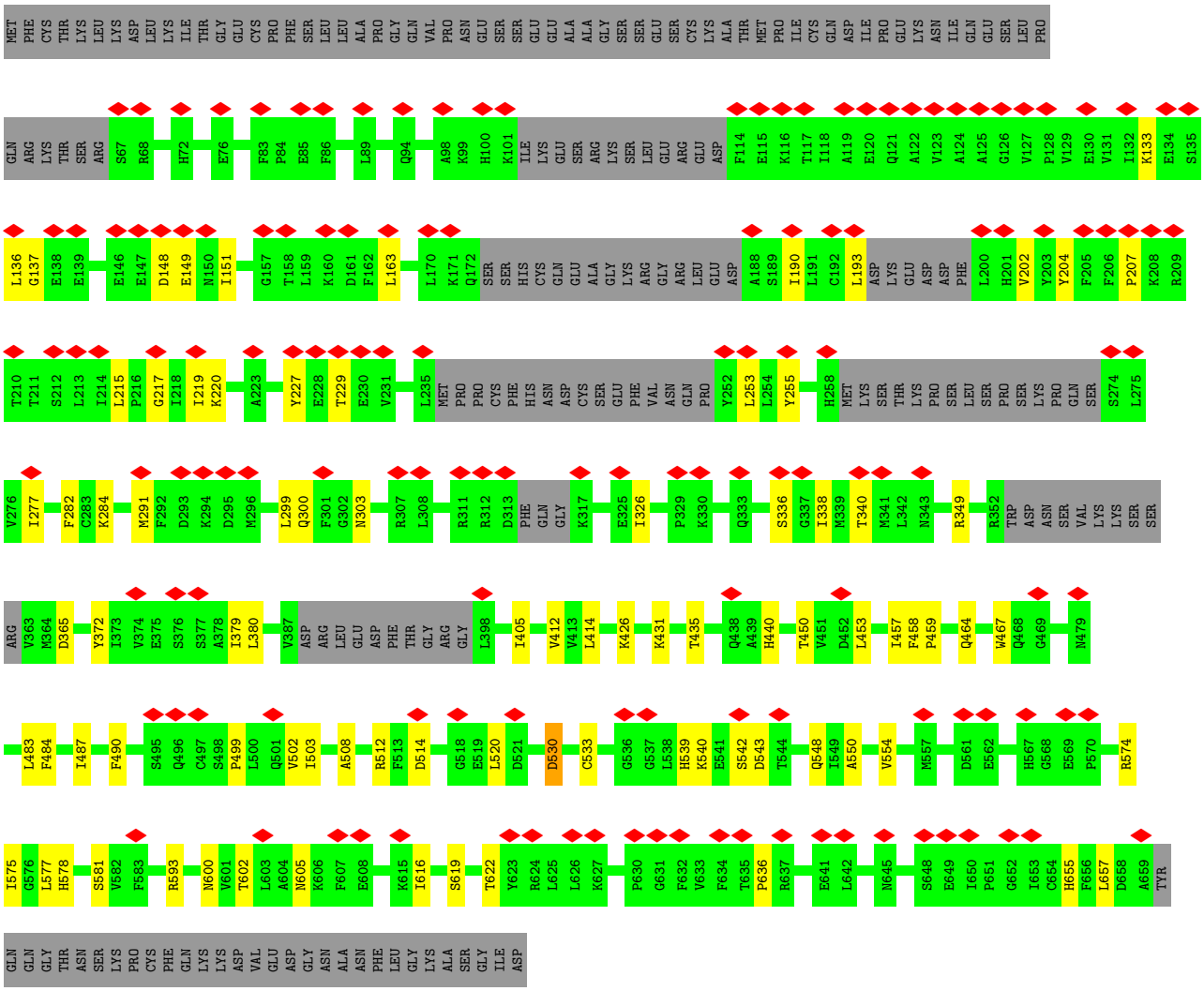
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
5	B	1	42	36	1	5	0

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.

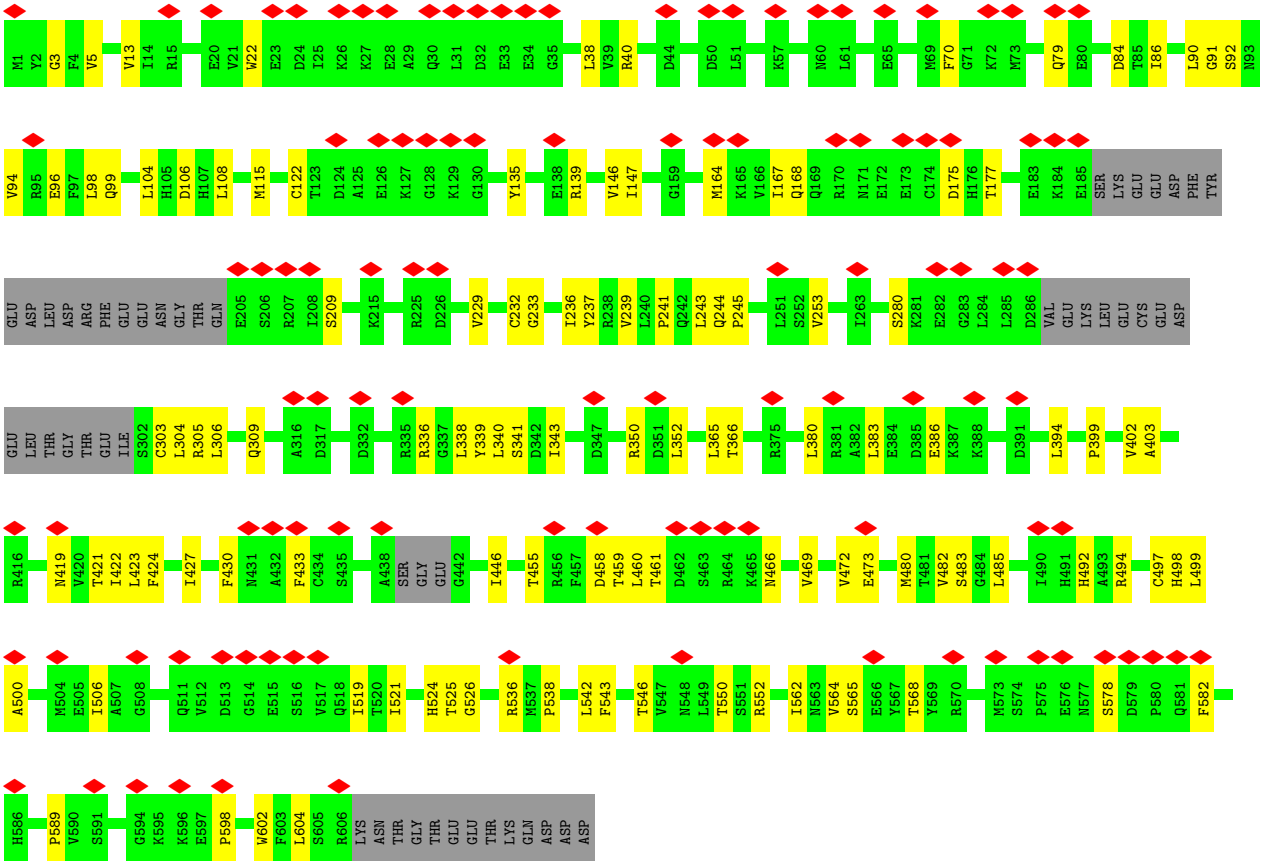
- Molecule 1: Guanylate cyclase soluble subunit alpha-1

Chain A:



- Molecule 2: Guanylate cyclase soluble subunit beta-1

Chain B:


M1 Y2 G3 F4 V5 V13 I14 R15 E20 V21 W22 E23 D24 I25 K26 K27 E28 A29 Q30 L31 D32 E33 E34 G35 L38 V39 R40 D44 D50 L51 K57 N60 L61 E65 M69 F70 G71 K72 M73 Q79 E80 D84 T85 I86 L90 G91 S92 N93
V94 R95 E96 F97 L98 Q99 L104 H105 D106 H107 L108 M115 C122 T123 D124 A125 E126 K127 G128 K129 G130 Y135 E138 R139 V146 I147 G159 M164 K165 V166 T167 Q168 Q169 R170 N171 E172 E173 C174 D175 H176 T177 E183 K184 E185 SER LYS GLU GLU ASP PHE TYR
GLU ASP LEU ASP ARG PHE GLU GLU ASN GLY THR GLN E205 S206 R207 I208 S209 K215 R225 D226 V229 C232 G233 I236 Y237 R238 V239 L240 P241 Q242 L243 Q244 P245 L251 S252 V253 I263 S280 K281 E282 G283 L284 L285 D286 VAL GLU LYS LEU GLU CYS GLU ASP
GLU LEU THR GLY THR GLU ILE S302 C303 L304 R305 L306 Q309 A316 D317 D332 R335 R336 G337 L338 Y339 L340 S341 D342 I343 D347 R350 D351 L352 L365 T366 R375 L380 R381 A382 L383 E384 D385 E386 K387 K388 D391 L394 P399 V402 A403
R416 N419 V420 T421 I422 L423 F424 I427 F430 N431 A432 F433 C434 S435 A438 SER GLY GLU G442 I446 T455 R456 F457 D458 T459 L460 T461 D462 S463 R464 K465 N466 V469 V472 E473 M480 T481 V482 S483 G484 L485 I490 H491 H492 A493 R494 C497 H498 L499
A500 M504 E505 I506 A507 G508 Q511 V512 D513 G514 E515 S516 V517 Q518 I519 T520 I521 H524 T525 G526 R536 M537 P538 L542 F543 T546 V547 N548 L549 T550 S551 R552 I562 N563 V564 S565 E566 Y567 T568 Y569 R570 M573 S574 P575 E576 N577 S578 D579 P580 Q581 F582
H586 P589 V590 S591 G594 K595 K596 E597 P598 W602 F603 L604 S605 R606 LYS ASN THR GLY THR GLU GLU THR LYS GLN ASP ASP ASP

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	453710	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	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 QUANTUM (4k x 4k)	Depositor
Maximum map value	0.159	Depositor
Minimum map value	-0.065	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.04	Depositor
Map size (Å)	250.79999, 250.79999, 250.79999	wwPDB
Map dimensions	240, 240, 240	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.045, 1.045, 1.045	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: Z90, MG, G2P

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| >5$	RMSZ	$\#\|Z\| >5$
1	A	0.26	0/3811	0.41	0/5174
2	B	0.26	0/4468	0.41	0/6068
All	All	0.26	0/8279	0.41	0/11242

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 (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	3742	0	3629	58	0
2	B	4372	0	4295	85	0
3	A	32	0	14	5	0
4	A	2	0	0	0	0
5	B	42	0	37	6	0
All	All	8190	0	7975	134	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 134 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:490:PHE:CD2	1:A:530:ASP:OD2	2.26	0.88
1:A:490:PHE:HD2	1:A:530:ASP:OD2	1.60	0.85
2:B:483:SER:HB2	2:B:492:HIS:HB2	1.71	0.72
1:A:520:LEU:O	1:A:548:GLN:NE2	2.24	0.70
2:B:526:GLY:HA3	2:B:546:THR:HG21	1.72	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	488/690 (71%)	473 (97%)	15 (3%)	0	100	100
2	B	563/619 (91%)	541 (96%)	21 (4%)	1 (0%)	47	79
All	All	1051/1309 (80%)	1014 (96%)	36 (3%)	1 (0%)	54	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	340	LEU

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	381/613 (62%)	380 (100%)	1 (0%)	92	96

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	466/555 (84%)	466 (100%)	0	100	100
All	All	847/1168 (72%)	846 (100%)	1 (0%)	93	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	530	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	408	HIS
2	B	451	ASN
2	B	466	ASN
1	A	422	GLN
2	B	99	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)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	G2P	A	1001	4	27,34,34	1.44	6 (22%)	33,54,54	1.98	8 (24%)
5	Z90	B	701	-	45,45,45	0.90	1 (2%)	58,58,58	0.97	2 (3%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	G2P	A	1001	4	-	8/15/38/38	0/3/3/3
5	Z90	B	701	-	-	13/30/30/30	0/4/4/4

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	701	Z90	CBO-CBN	5.15	1.48	1.40
3	A	1001	G2P	C5-C6	4.02	1.48	1.41
3	A	1001	G2P	PA-O5'	2.58	1.61	1.57
3	A	1001	G2P	C5-C4	2.35	1.47	1.40
3	A	1001	G2P	PB-O3B	2.31	1.60	1.58

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1001	G2P	C2-N3-C4	4.85	120.90	115.36
3	A	1001	G2P	C5-C6-N1	-4.05	117.89	123.43
3	A	1001	G2P	PB-O3B-PG	-4.05	118.35	132.62
3	A	1001	G2P	C2-N1-C6	4.03	122.33	115.93
3	A	1001	G2P	C4-C5-C6	-3.64	117.32	120.80

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001	G2P	PA-C3A-PB-O3B
3	A	1001	G2P	PA-C3A-PB-O1B
3	A	1001	G2P	PA-C3A-PB-O2B
3	A	1001	G2P	PB-C3A-PA-O1A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	1001	G2P	PB-C3A-PA-O2A

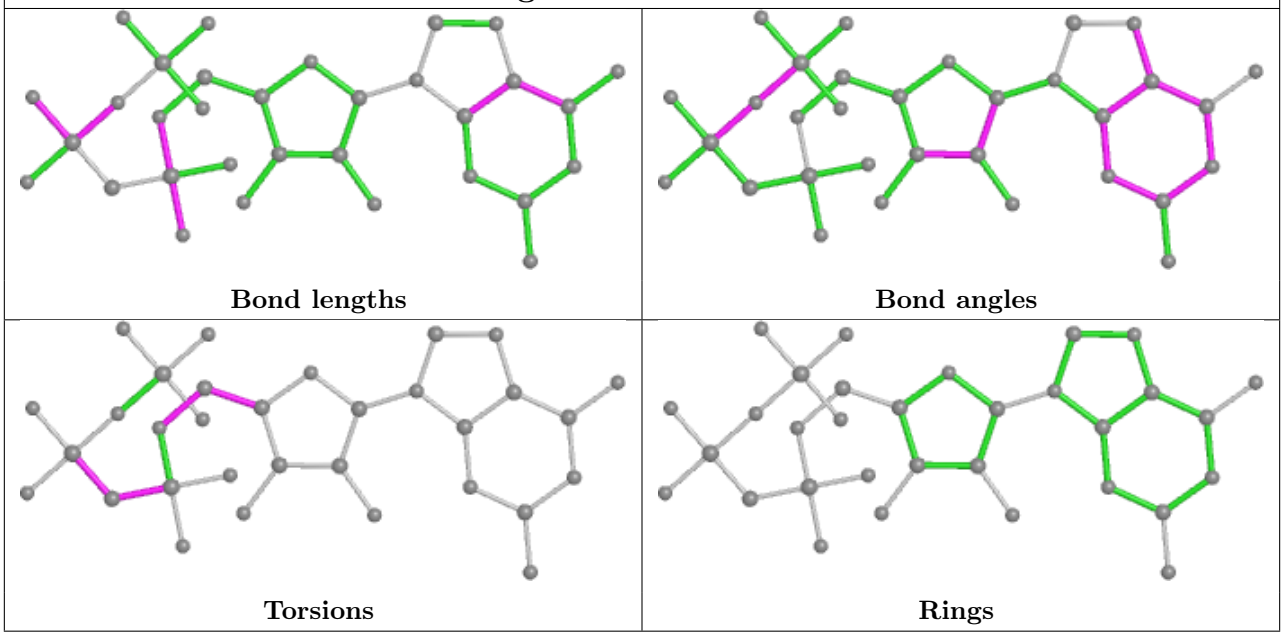
There are no ring outliers.

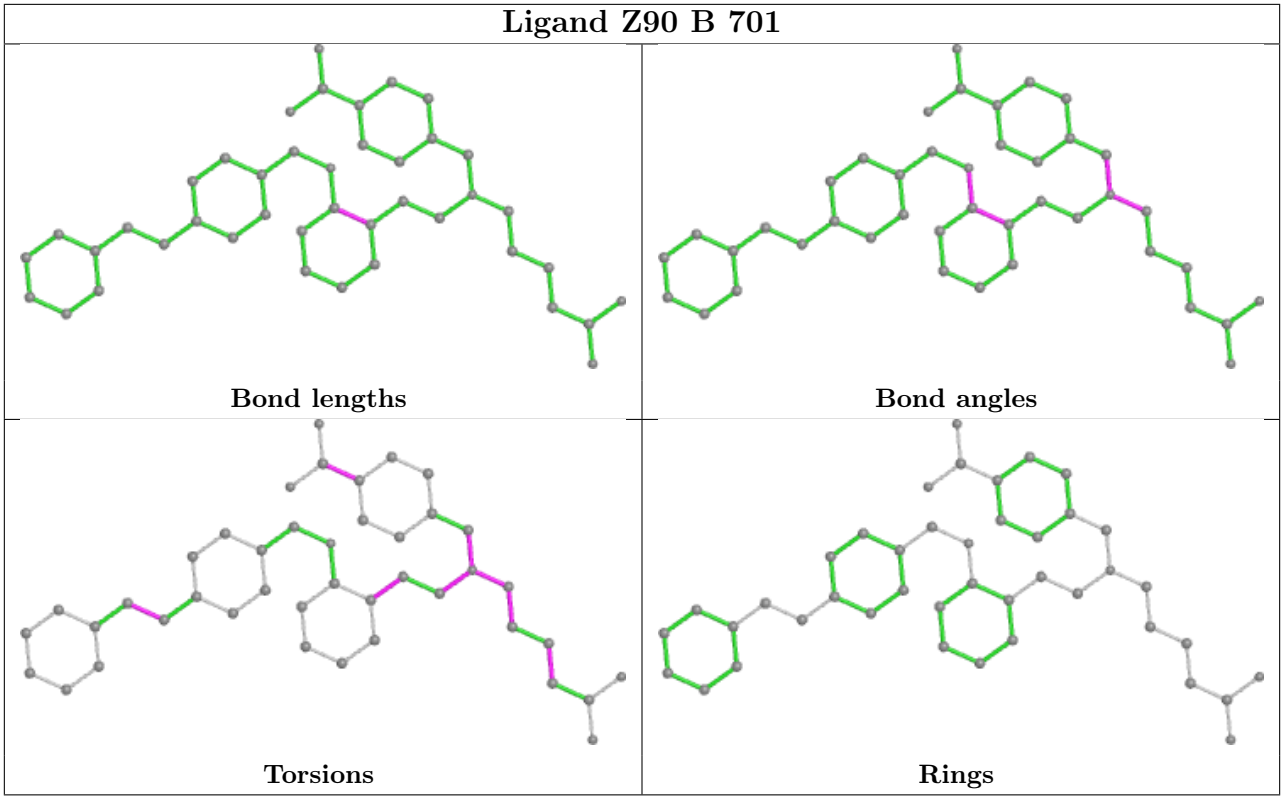
2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	G2P	5	0
5	B	701	Z90	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand G2P A 1001





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.

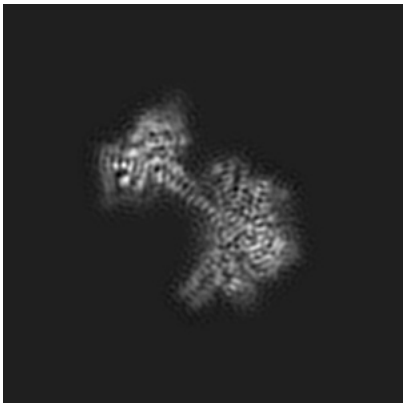
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-30621. These allow visual inspection of the internal detail of the map and identification of artifacts.

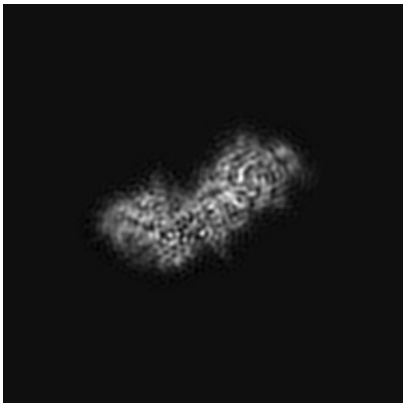
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

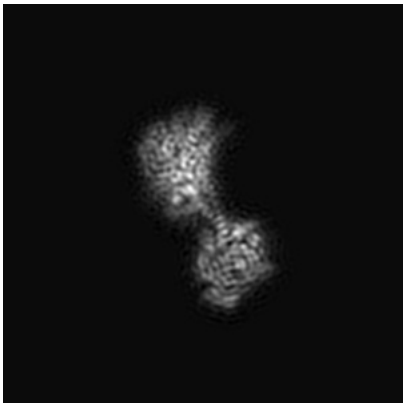
6.1.1 Primary map



X



Y

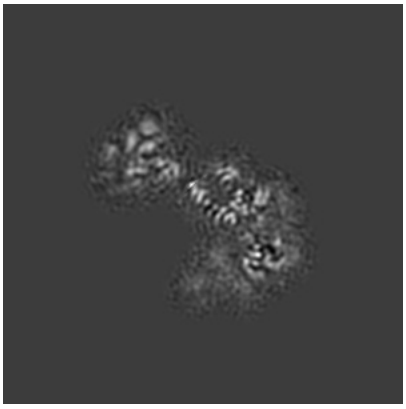


Z

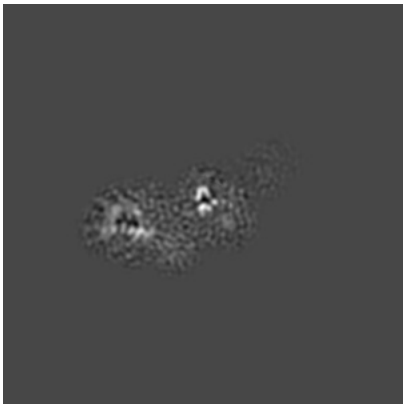
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

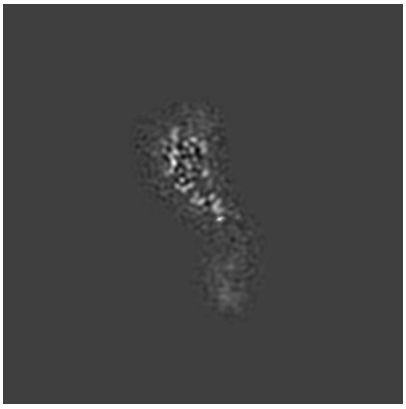
6.2.1 Primary map



X Index: 120



Y Index: 120

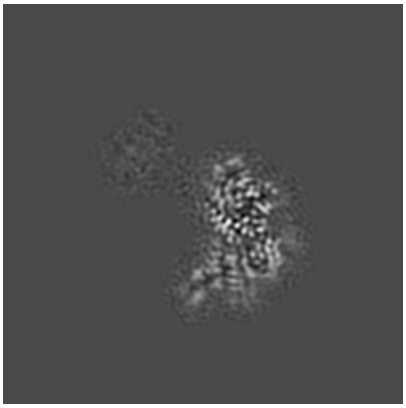


Z Index: 120

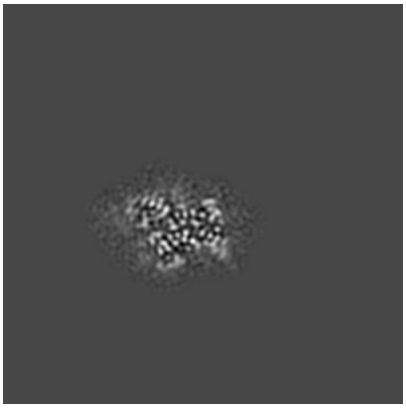
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

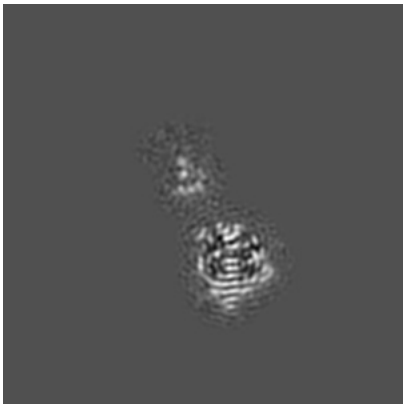
6.3.1 Primary map



X Index: 112



Y Index: 152



Z Index: 142

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

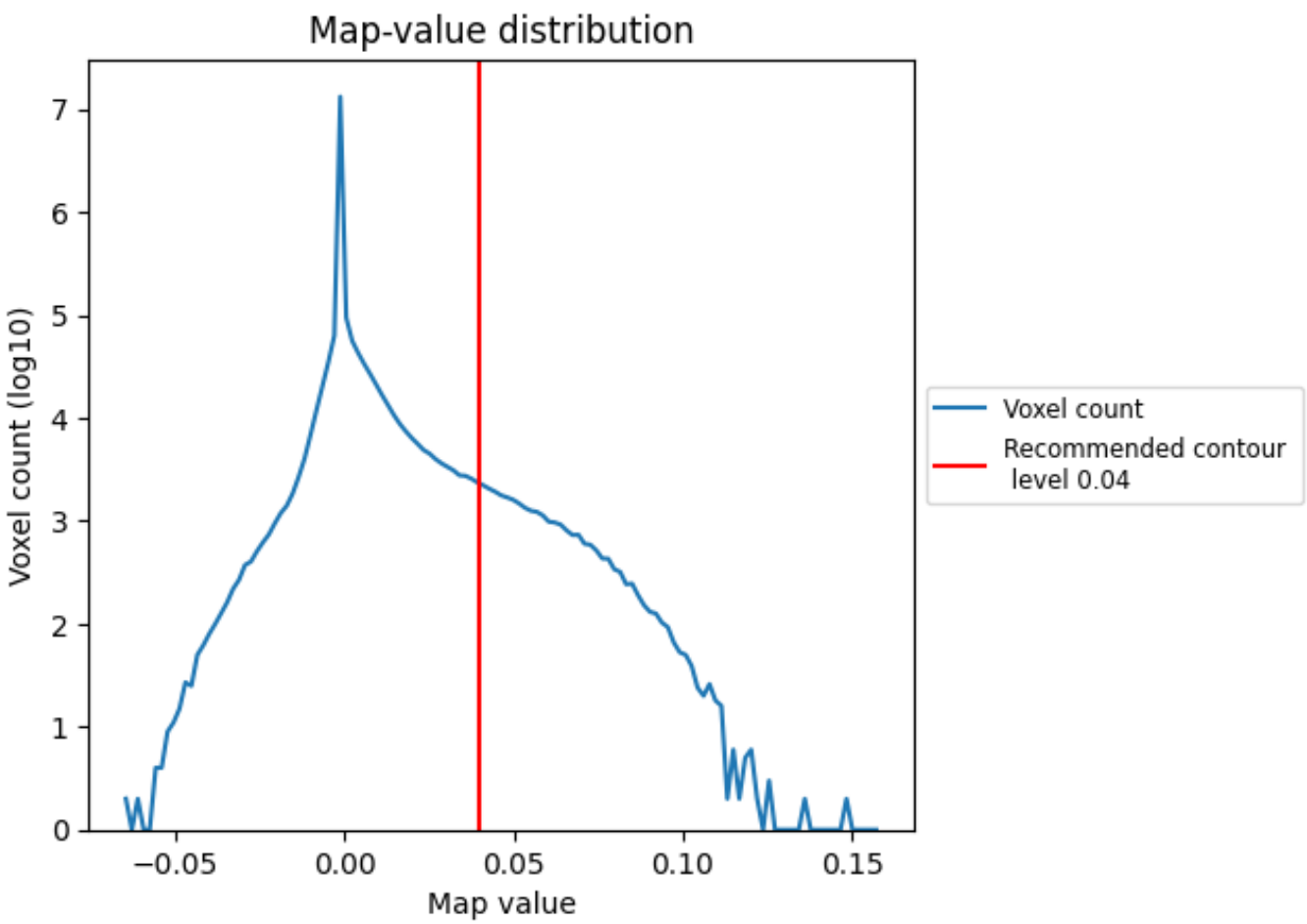
6.5 Mask visualisation (i)

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

7 Map analysis (i)

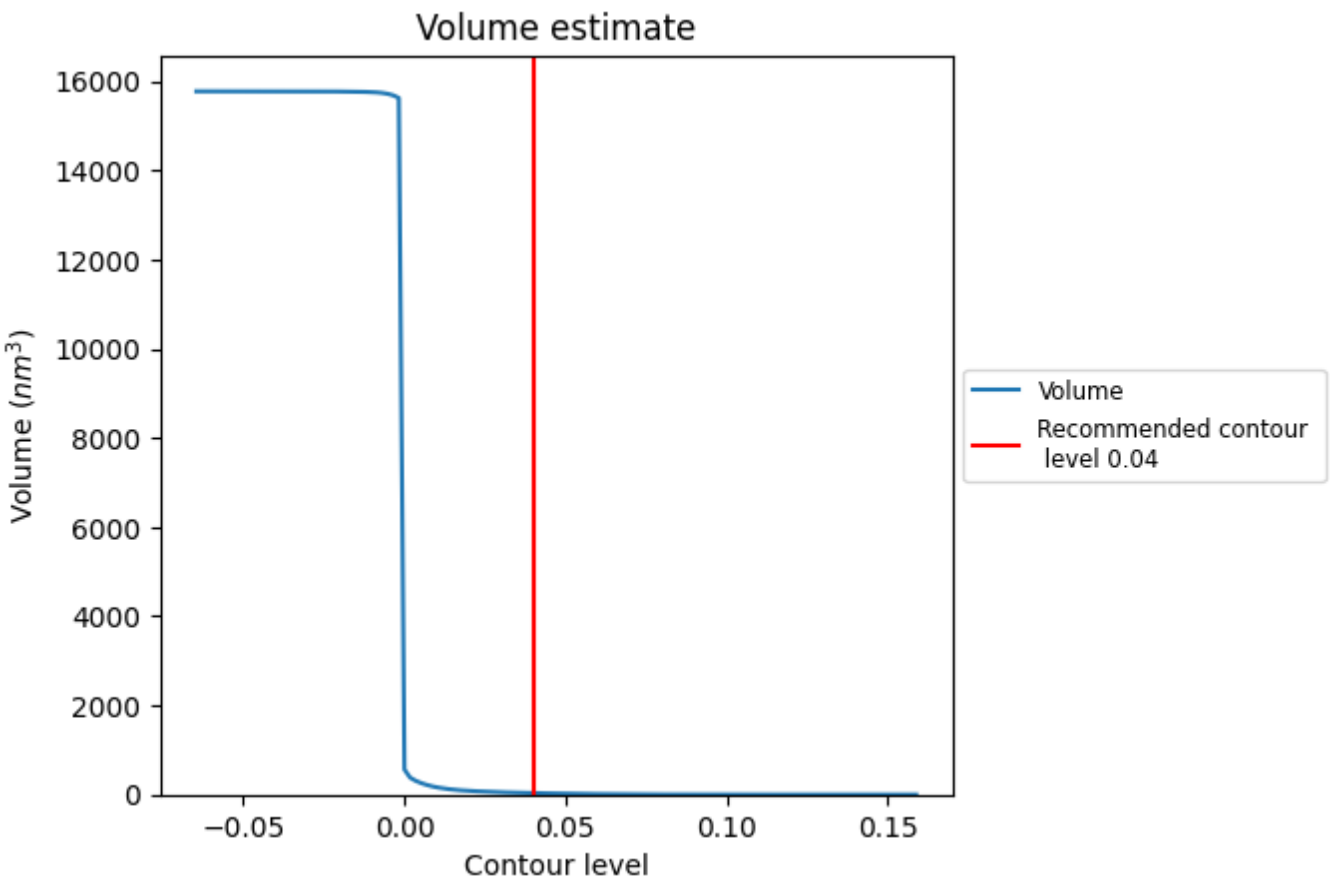
This section contains the results of statistical analysis of the map.

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.

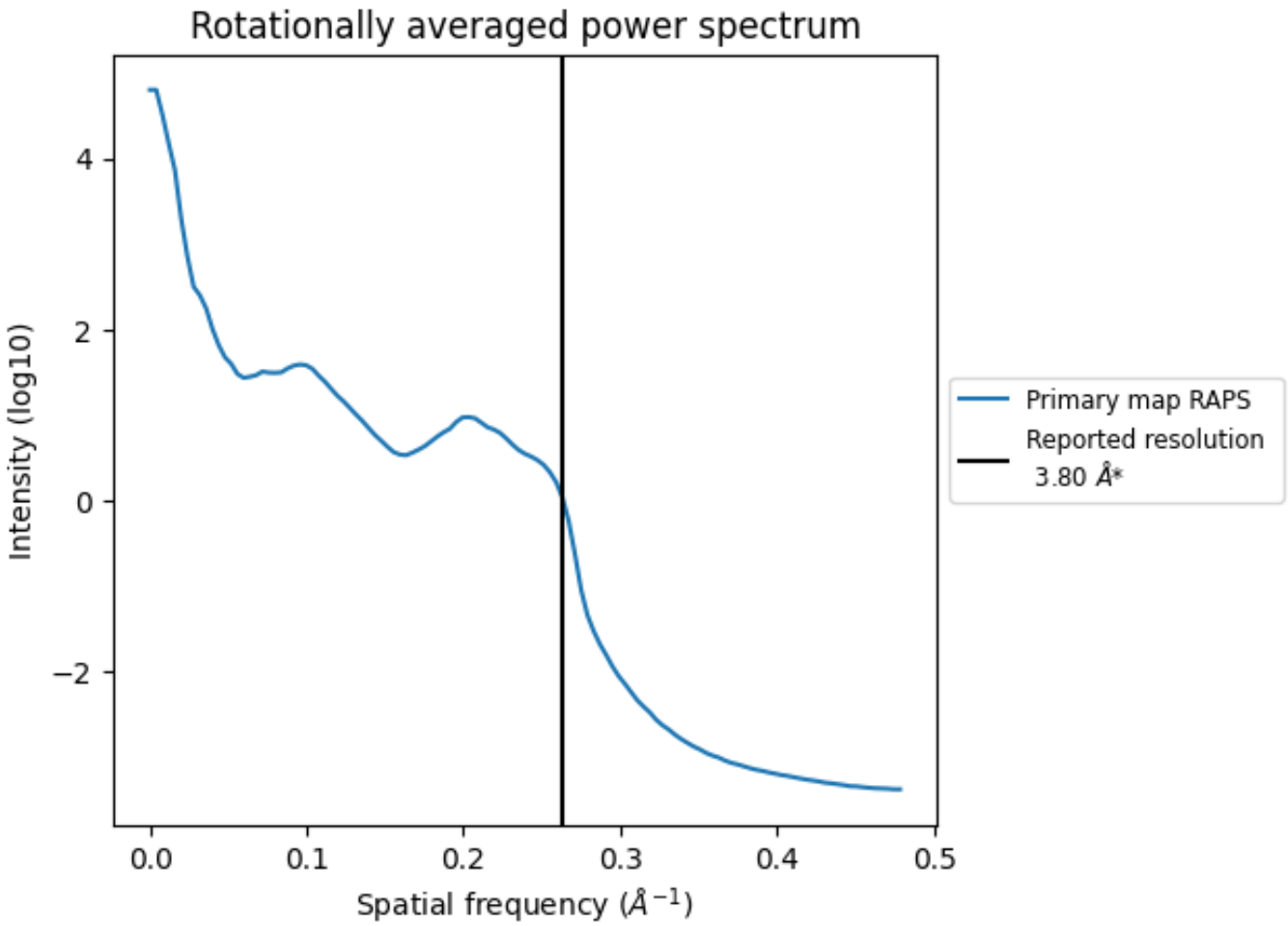
7.2 Volume estimate (i)



The volume at the recommended contour level is 33 nm^3; this corresponds to an approximate mass of 30 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 0.263 Å^{-1}

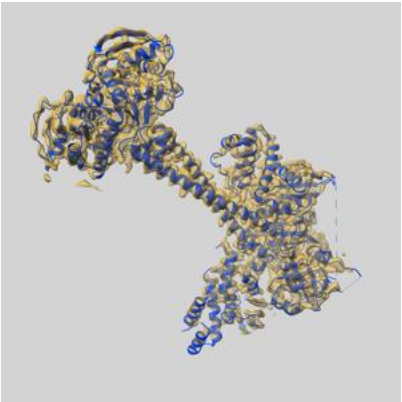
8 Fourier-Shell correlation (i)

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

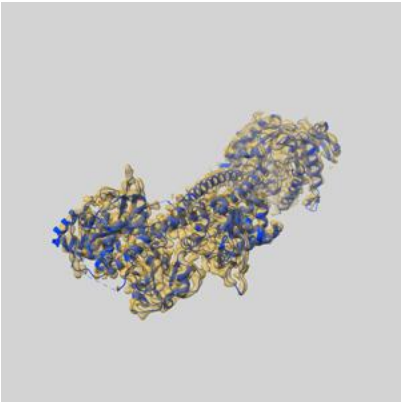
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-30621 and PDB model 7D9U. Per-residue inclusion information can be found in section 3 on page 5.

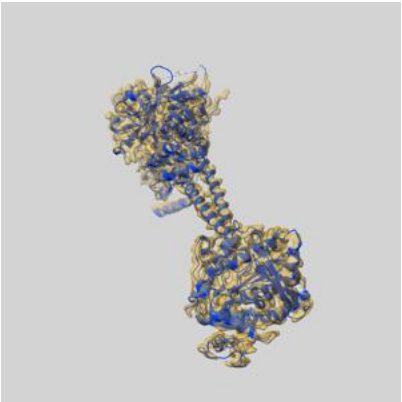
9.1 Map-model overlay (i)



X



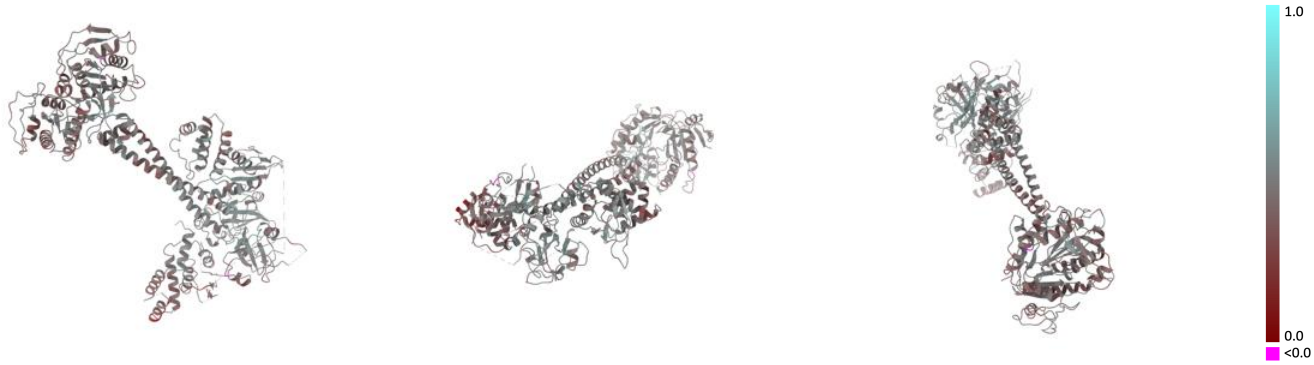
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.04 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.

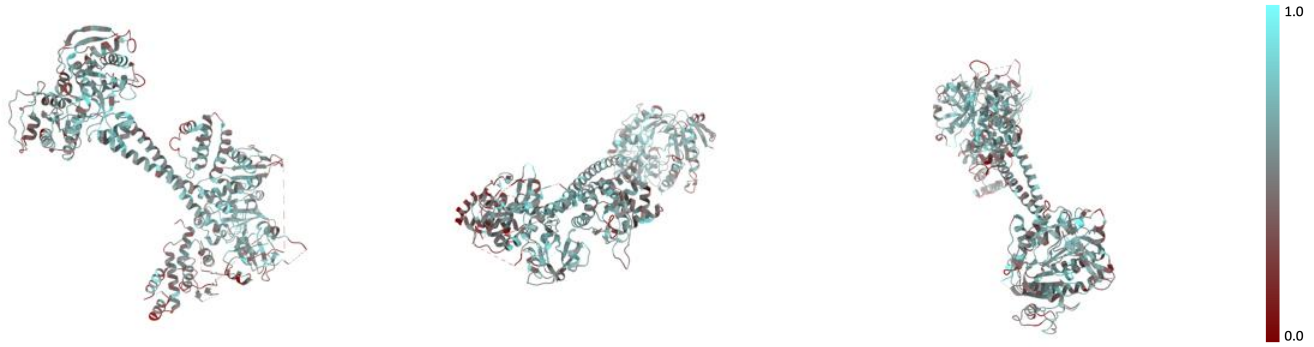
9.2 Q-score mapped to coordinate model (i)



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.

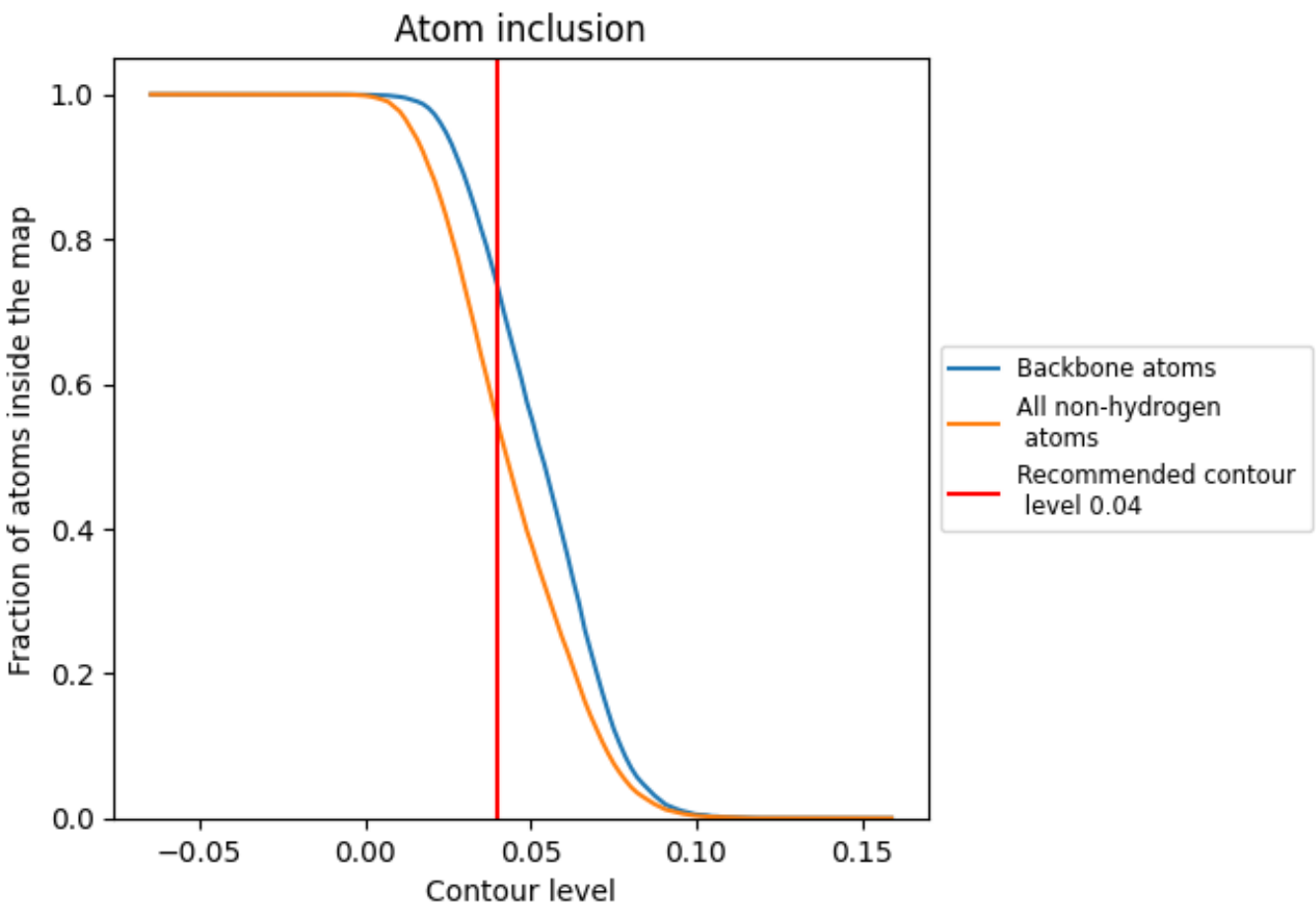
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.04).

9.4 Atom inclusion (i)



At the recommended contour level, 73% of all backbone atoms, 55% 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 (0.04) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5462	0.4450
A	0.5165	0.4280
B	0.5717	0.4590