wwPDB EM Validation Summary Report

Mar 19, 2026 – 08:19 PM UTC

PDB ID	:	9GY0 / pdb_00009gy0
EMDB ID	:	EMD-51680
Title	:	Cryo_EM structure of human FAN1 R507H mutant in complex with 5' flap DNA substrate and PCNA
Authors	:	Jeyasankar, G.; Salerno-Kochan, A.; Thomsen, M.
Deposited on	:	2024-10-01
Resolution	:	3.42 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:

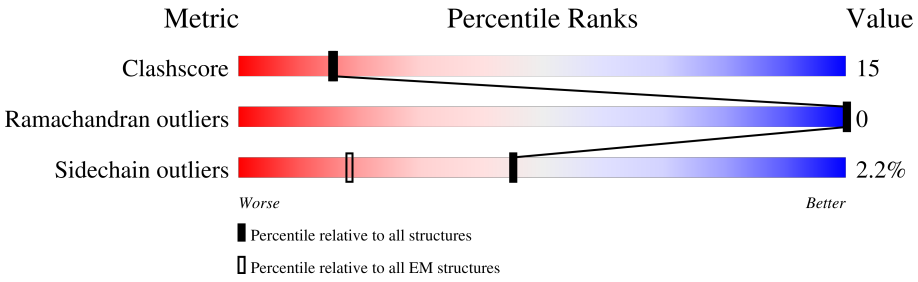
EMDB validation analysis	:	0.0.1.dev132
MolProbity	:	4-5-2 with Phenix2.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics	:	NOT EXECUTED
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 3.42 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	229148	23984
Ramachandran outliers	224038	23583
Sidechain outliers	223484	23102

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%

Mol	Chain	Length	Quality of chain
1	A	1021	42% 16% 42%
2	B	40	38% 38% 25%
3	C	28	18% 64% 18%
4	D	21	14% 43% 43%
5	E	263	66% 31% .
5	F	263	59% 29% . 11%
5	G	263	58% 37% . .

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 11866 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 Fanconi-associated nuclease 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	596	4816	3072	860	860	24	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q9Y2M0
A	-2	PRO	-	expression tag	UNP Q9Y2M0
A	-1	MET	-	expression tag	UNP Q9Y2M0
A	0	GLY	-	expression tag	UNP Q9Y2M0
A	507	HIS	ARG	engineered mutation	UNP Q9Y2M0

- Molecule 2 is a DNA chain called DNA (Continuous).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	B	30	605	287	103	185	30	0	0

- Molecule 3 is a DNA chain called DNA (Pre-Nick).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	C	23	471	224	88	137	22	0	0

- Molecule 4 is a DNA chain called DNA (Post-Nick).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	D	12	251	117	51	71	12	0	0

- Molecule 5 is a protein called Proliferating cell nuclear antigen.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	256	Total	C	N	O	S	0	1
			1964	1233	323	392	16		
5	F	235	Total	C	N	O	S	0	0
			1812	1143	298	355	16		
5	G	253	Total	C	N	O	S	0	0
			1946	1221	319	390	16		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	GLY	-	expression tag	UNP P12004
E	0	PRO	-	expression tag	UNP P12004
F	-1	GLY	-	expression tag	UNP P12004
F	0	PRO	-	expression tag	UNP P12004
G	-1	GLY	-	expression tag	UNP P12004
G	0	PRO	-	expression tag	UNP P12004

- Molecule 6 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

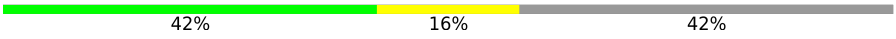
Mol	Chain	Residues	Atoms		AltConf
6	A	1	Total	Ca	0
			1	1	

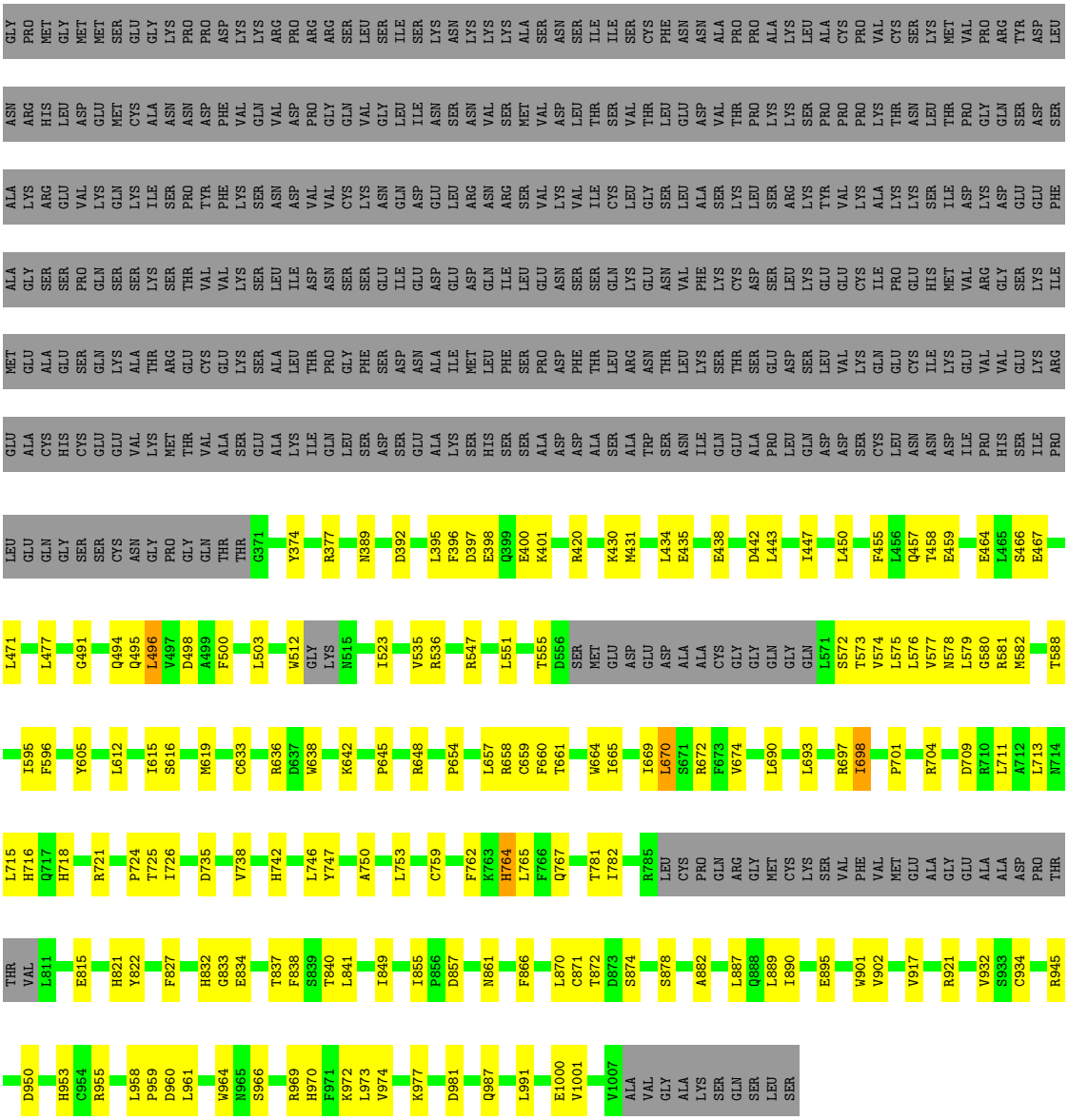
3 Residue-property plots (i)

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. 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. 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: Fanconi-associated nuclease 1

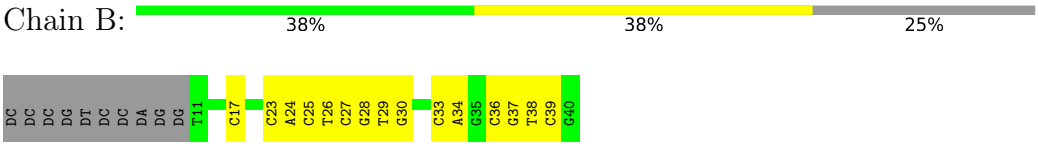
Chain A:





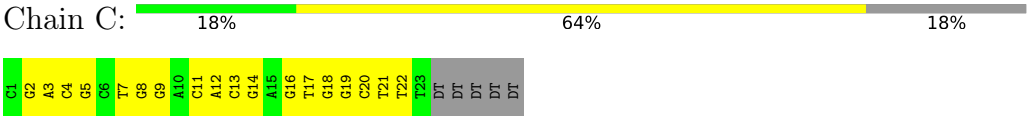
- Molecule 2: DNA (Continuous)

Chain B:



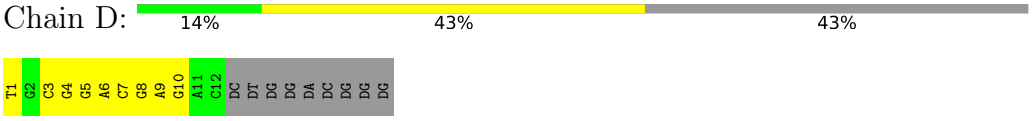
• Molecule 3: DNA (Pre-Nick)

Chain C:



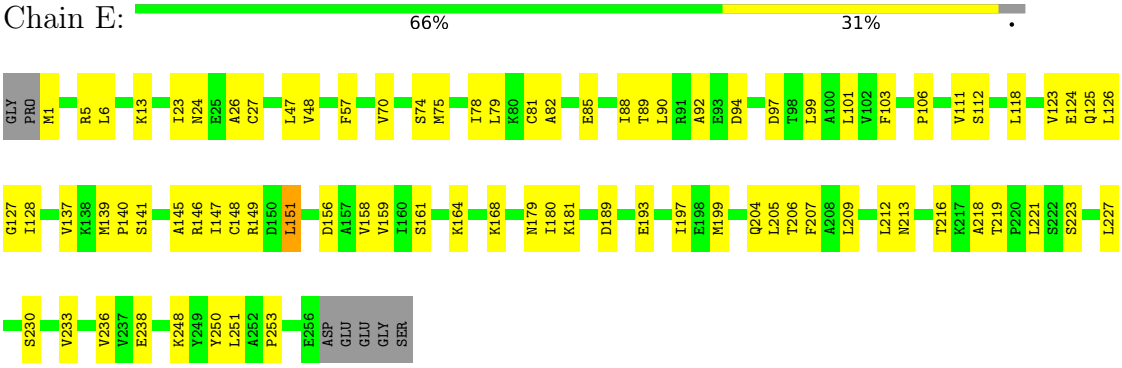
• Molecule 4: DNA (Post-Nick)

Chain D:



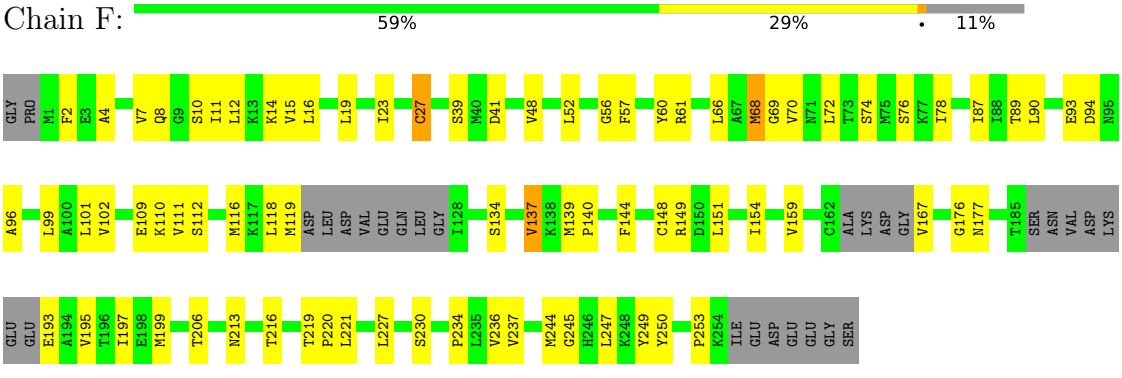
• Molecule 5: Proliferating cell nuclear antigen

Chain E:



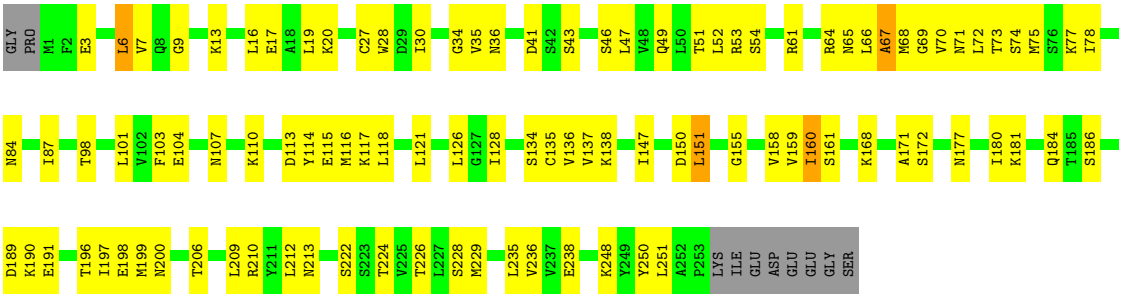
• Molecule 5: Proliferating cell nuclear antigen

Chain F:



• Molecule 5: Proliferating cell nuclear antigen

Chain G:


GLY PRO M1 F2 E3 L6 V7 Q8 G9 K13 L16 E17 A18 L19 K20 C27 W28 D29 I30 G34 V35 N36 D41 S42 S43 S46 L47 V48 Q49 L50 T51 L52 R53 S54 R61 R64 N65 L66 A67 M68 G69 V70 N71 L72 T73 S74 M75 S76 K77 I78
N84 I87 T98 L101 V102 F103 E104 N107 K110 D113 Y114 E115 M116 K117 L118 L121 L126 G127 I128 S134 C135 V136 V137 K138 I147 D150 L151 G155 V158 V159 I160 S161 K168 A171 S172 N177 I180 K181 Q184 T185 S186
D189 K190 E191 T196 I197 E198 M199 N200 T206 L209 R210 Y211 L212 N213 S222 S223 T224 V225 T226 L227 S228 M229 L235 V236 V237 E238 K248 Y249 Y250 L251 A252 P253 LYS ILE GLU ASP GLU GLU GLY SER

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	487606	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS GLACIOS	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	60	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.21	0/4920	0.36	0/6653
2	B	0.15	0/674	0.32	0/1036
3	C	0.23	0/528	0.42	0/814
4	D	0.13	0/282	0.27	0/432
5	E	0.22	0/1990	0.38	0/2689
5	F	0.21	1/1835 (0.1%)	0.34	0/2475
5	G	0.28	0/1972	0.52	2/2665 (0.1%)
All	All	0.22	1/12201 (0.0%)	0.39	2/16764 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	68	MET	C-N	5.38	1.41	1.33

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	G	67	ALA	CA-C-N	-5.77	114.17	122.36
5	G	67	ALA	C-N-CA	-5.77	114.17	122.36

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	4816	0	4852	129	0
2	B	605	0	338	11	0
3	C	471	0	260	20	0
4	D	251	0	134	9	0
5	E	1964	0	1973	59	0
5	F	1812	0	1829	55	0
5	G	1946	0	1947	78	0
6	A	1	0	0	0	0
All	All	11866	0	11333	338	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 15.

The worst 5 of 338 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)
5:E:1:MET:HA	5:E:92:ALA:O	1.59	1.02
5:G:16:LEU:HD11	5:G:75:MET:HB3	1.46	0.95
1:A:458:THR:HG22	1:A:459:GLU:H	1.34	0.90
1:A:458:THR:HG22	1:A:459:GLU:N	1.96	0.81
1:A:458:THR:HG22	1:A:459:GLU:OE1	1.83	0.79

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	588/1021 (58%)	570 (97%)	18 (3%)	0	100	100
5	E	254/263 (97%)	242 (95%)	12 (5%)	0	100	100
5	F	227/263 (86%)	221 (97%)	6 (3%)	0	100	100
5	G	251/263 (95%)	237 (94%)	14 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1320/1810 (73%)	1270 (96%)	50 (4%)	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	525/902 (58%)	515 (98%)	10 (2%)	50 66
5	E	223/229 (97%)	218 (98%)	5 (2%)	45 63
5	F	206/229 (90%)	200 (97%)	6 (3%)	37 59
5	G	221/229 (96%)	216 (98%)	5 (2%)	44 62
All	All	1175/1589 (74%)	1149 (98%)	26 (2%)	45 63

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	156	ASP
5	F	109	GLU
5	G	160	ILE
5	F	27	CYS
5	F	137	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
5	G	184	GLN
5	G	213	ASN
5	E	95	ASN
5	F	153	HIS
5	F	179	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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-51680. 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

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

6.5 Orthogonal surface views

This section was not generated.

6.6 Mask visualisation

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

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

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

9 Map-model fit

This section was not generated.