wwPDB EM Validation Summary Report (i)

Mar 10, 2024 – 10:26 PM EDT

PDB ID	:	6OKR
EMDB ID	:	EMD-20102
Title	:	CDTb Pre-Insertion form Modeled from Cryo-EM Map Reconstructed using C7 Symmetry
Authors	:	Lacy, D.B.; Sheedlo, M.J.; Anderson, D.M.
Deposited on	:	2019-04-15
Resolution	:	4.20 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:

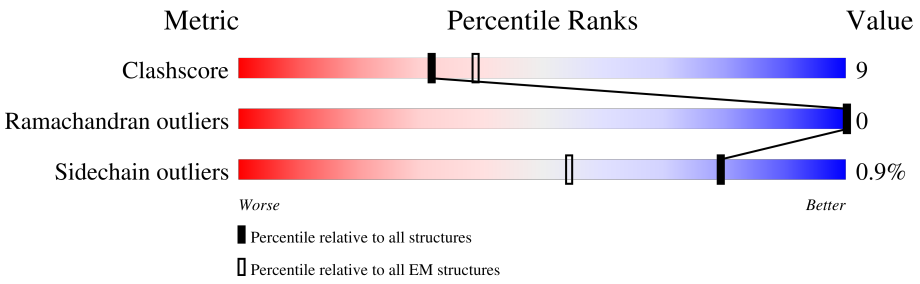
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 4.20 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	876	5% 30% 9% 61%
1	B	876	5% 30% 9% 61%
1	C	876	5% 30% 9% 61%
1	D	876	5% 30% 8% 61%
1	E	876	5% 30% 9% 61%
1	F	876	5% 30% 9% 61%
1	G	876	5% 30% 8% 61%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 18613 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 ADP-ribosyltransferase binding component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	342	2659	1660	433	561	5	0	0
1	B	342	2659	1660	433	561	5	0	0
1	C	342	2659	1660	433	561	5	0	0
1	D	342	2659	1660	433	561	5	0	0
1	E	342	2659	1660	433	561	5	0	0
1	F	342	2659	1660	433	561	5	0	0
1	G	342	2659	1660	433	561	5	0	0

There are 14 discrepancies between the modelled and reference sequences:

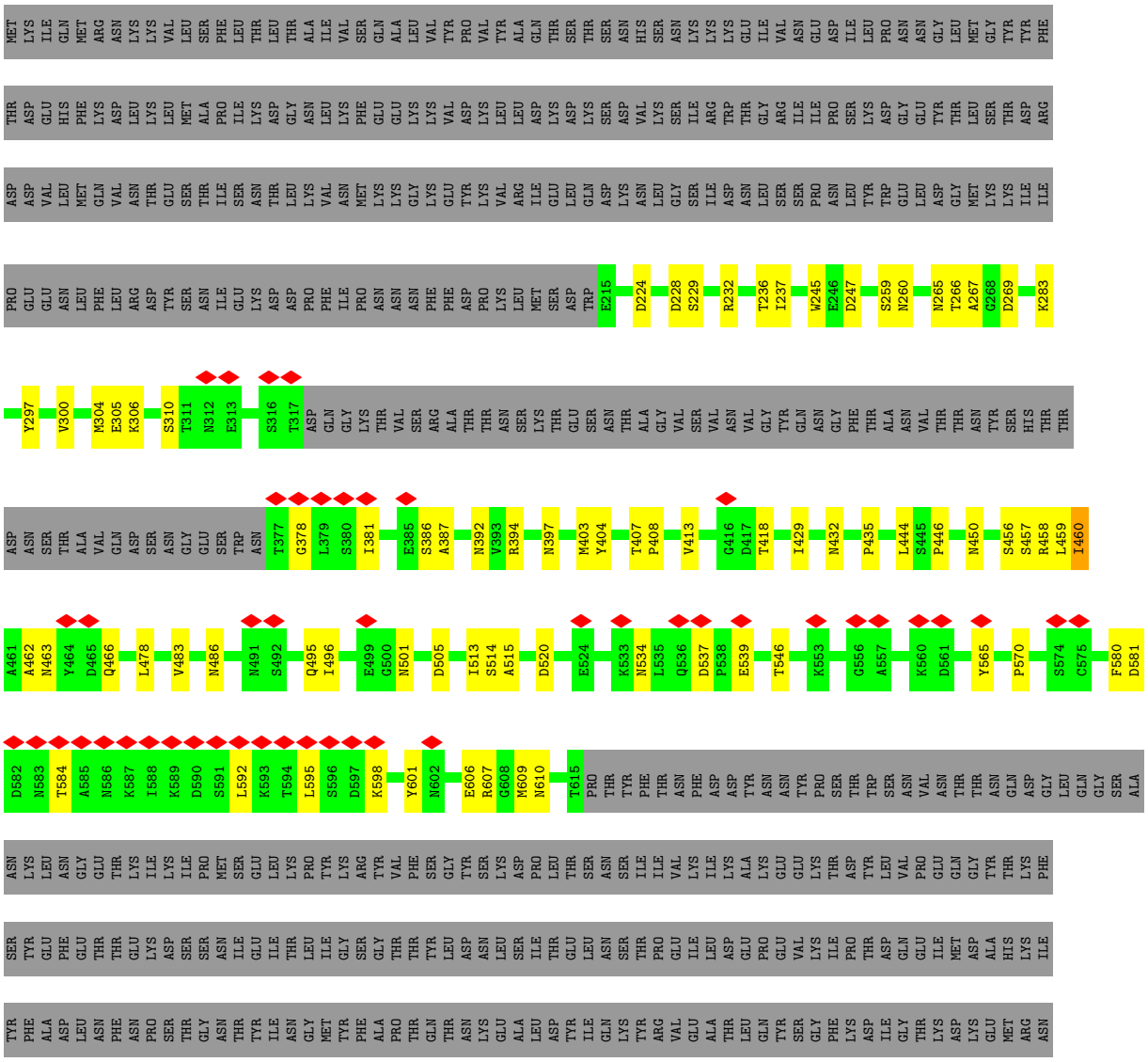
Chain	Residue	Modelled	Actual	Comment	Reference
A	461	ALA	PRO	conflict	UNP A8DS70
A	462	ALA	ILE	conflict	UNP A8DS70
B	461	ALA	PRO	conflict	UNP A8DS70
B	462	ALA	ILE	conflict	UNP A8DS70
C	461	ALA	PRO	conflict	UNP A8DS70
C	462	ALA	ILE	conflict	UNP A8DS70
D	461	ALA	PRO	conflict	UNP A8DS70
D	462	ALA	ILE	conflict	UNP A8DS70
E	461	ALA	PRO	conflict	UNP A8DS70
E	462	ALA	ILE	conflict	UNP A8DS70
F	461	ALA	PRO	conflict	UNP A8DS70
F	462	ALA	ILE	conflict	UNP A8DS70
G	461	ALA	PRO	conflict	UNP A8DS70
G	462	ALA	ILE	conflict	UNP A8DS70

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.

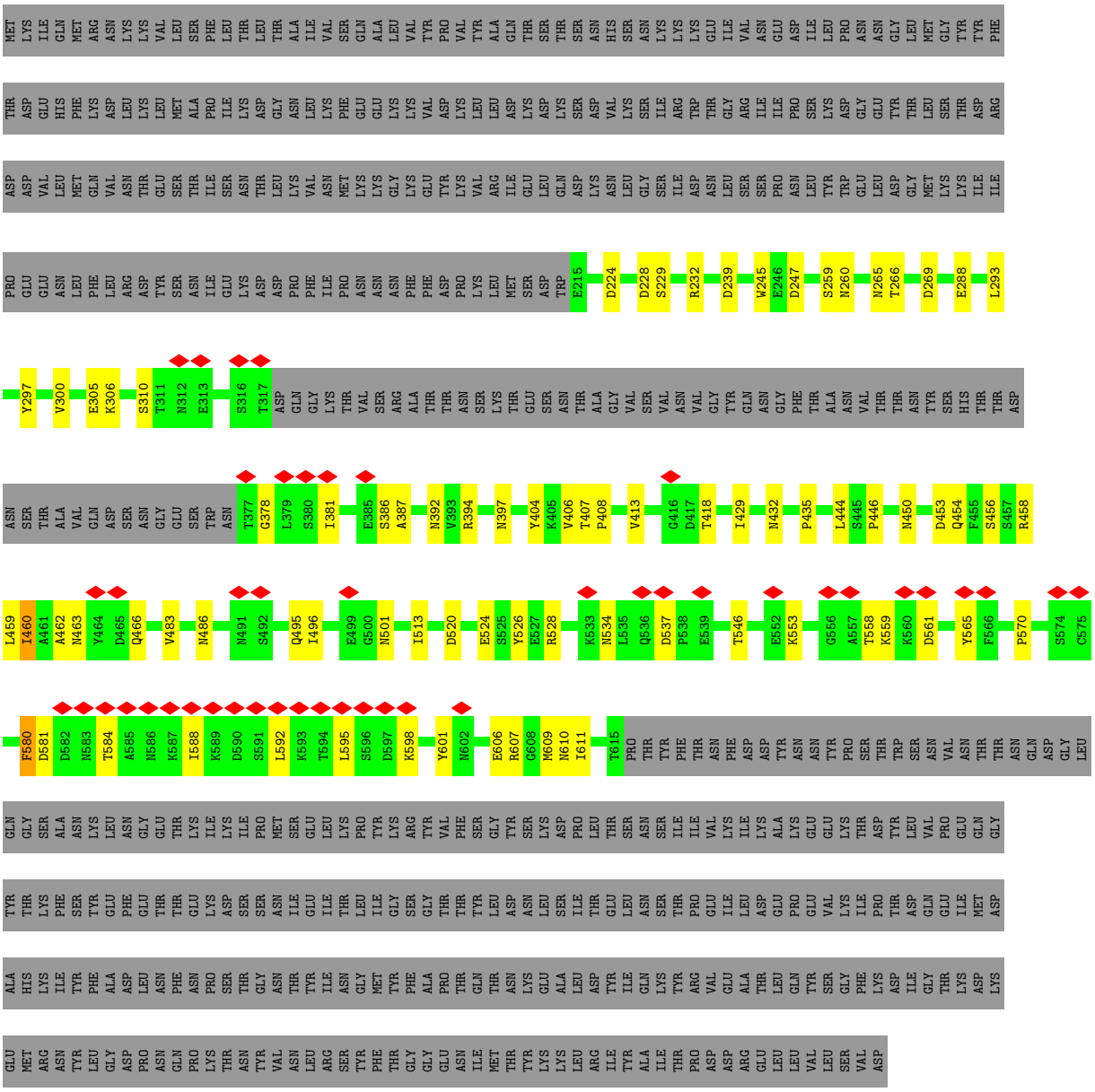
- Molecule 1: ADP-ribosyltransferase binding component

Chain A: 5% 30% 9% 61%



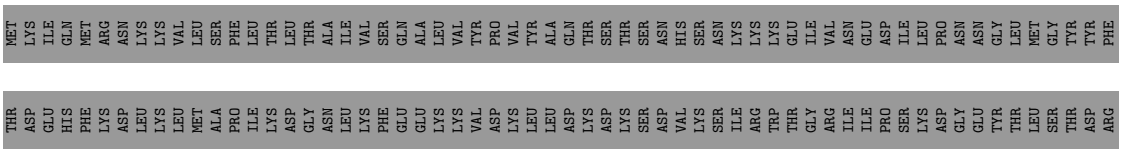
• Molecule 1: ADP-ribosyltransferase binding component

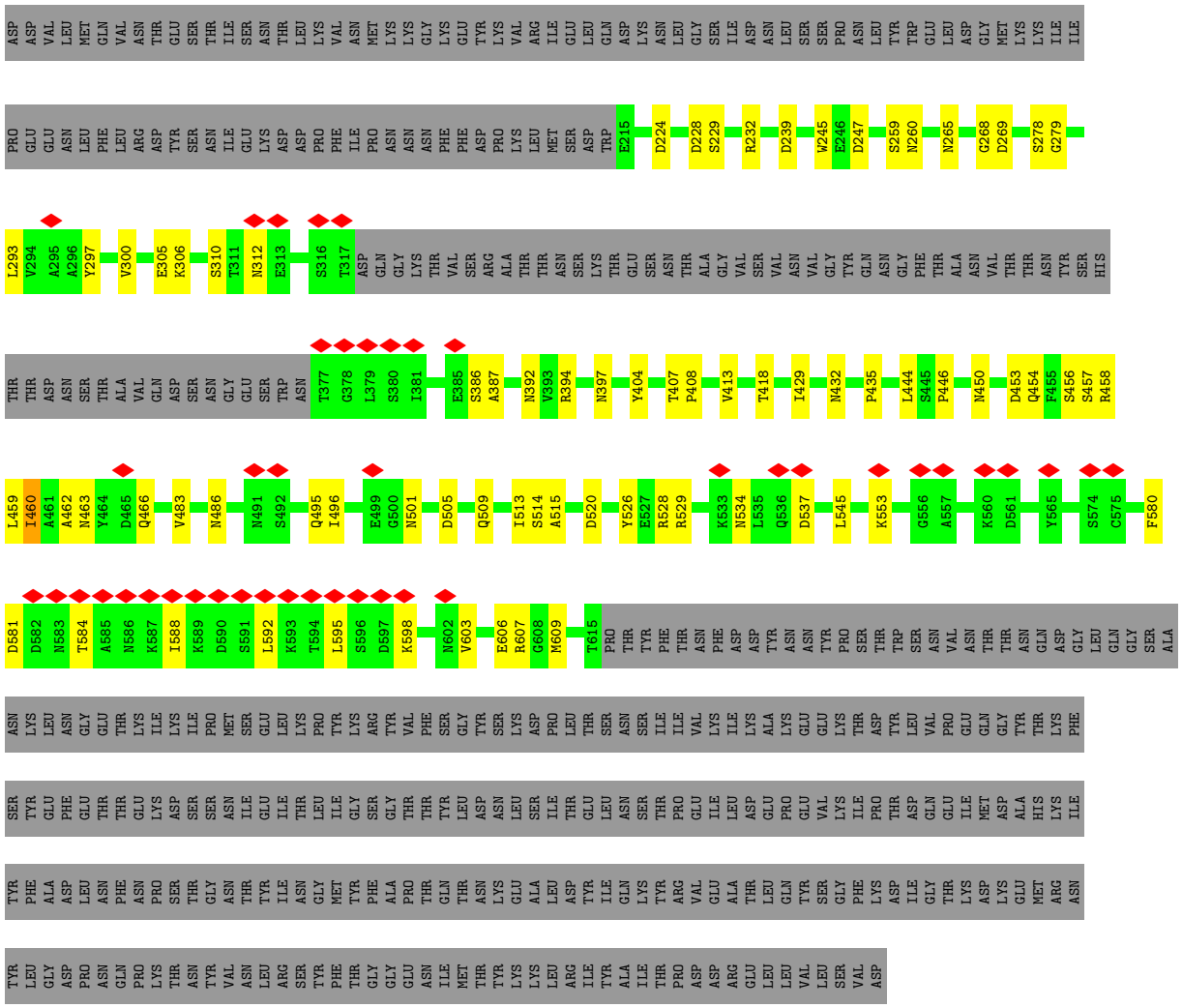
Chain B:



• Molecule 1: ADP-ribosyltransferase binding component

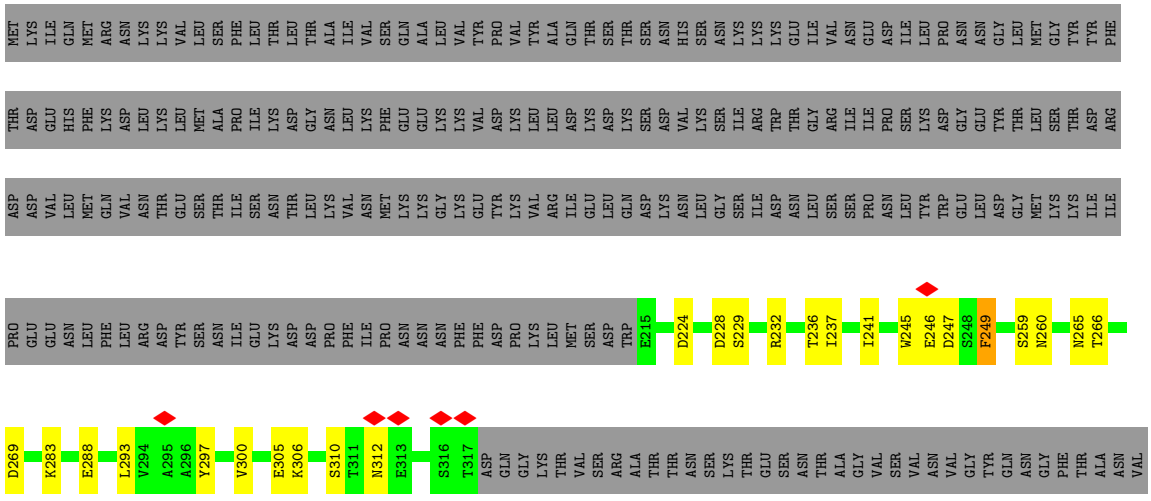
Chain C:

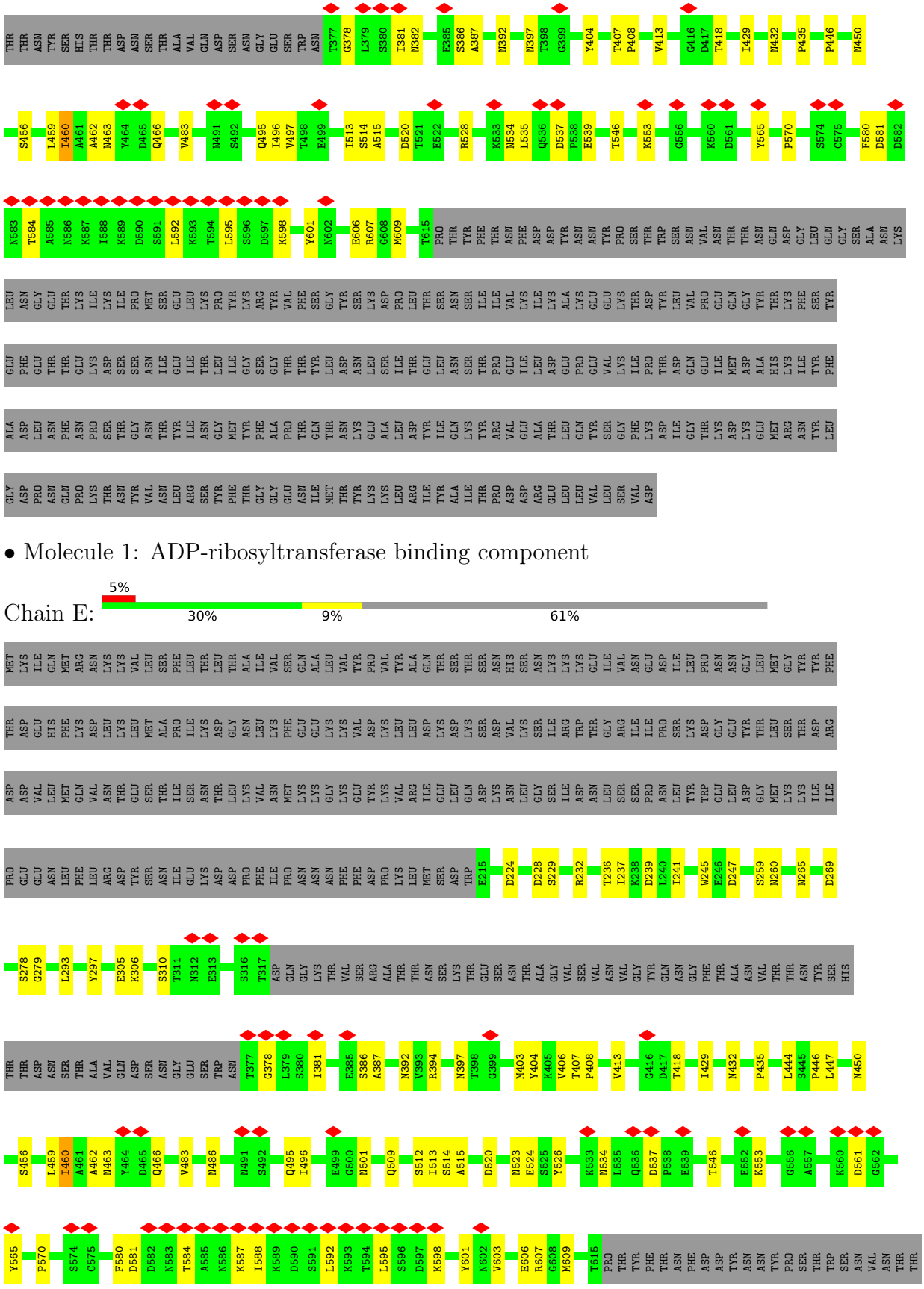




- Molecule 1: ADP-ribosyltransferase binding component

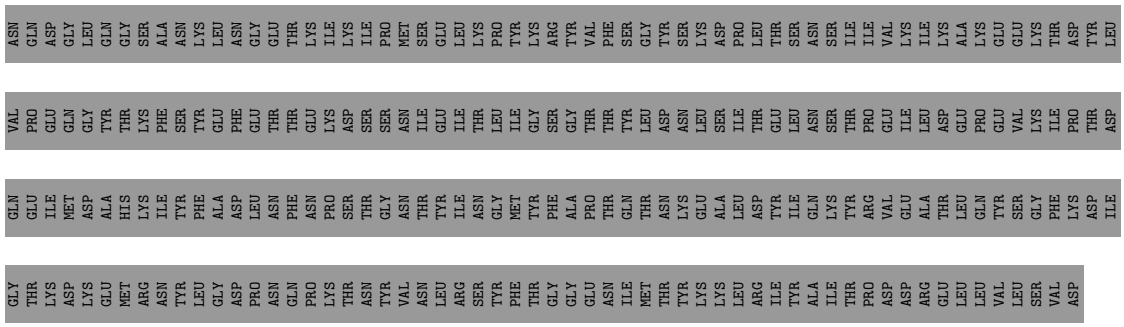
Chain D:



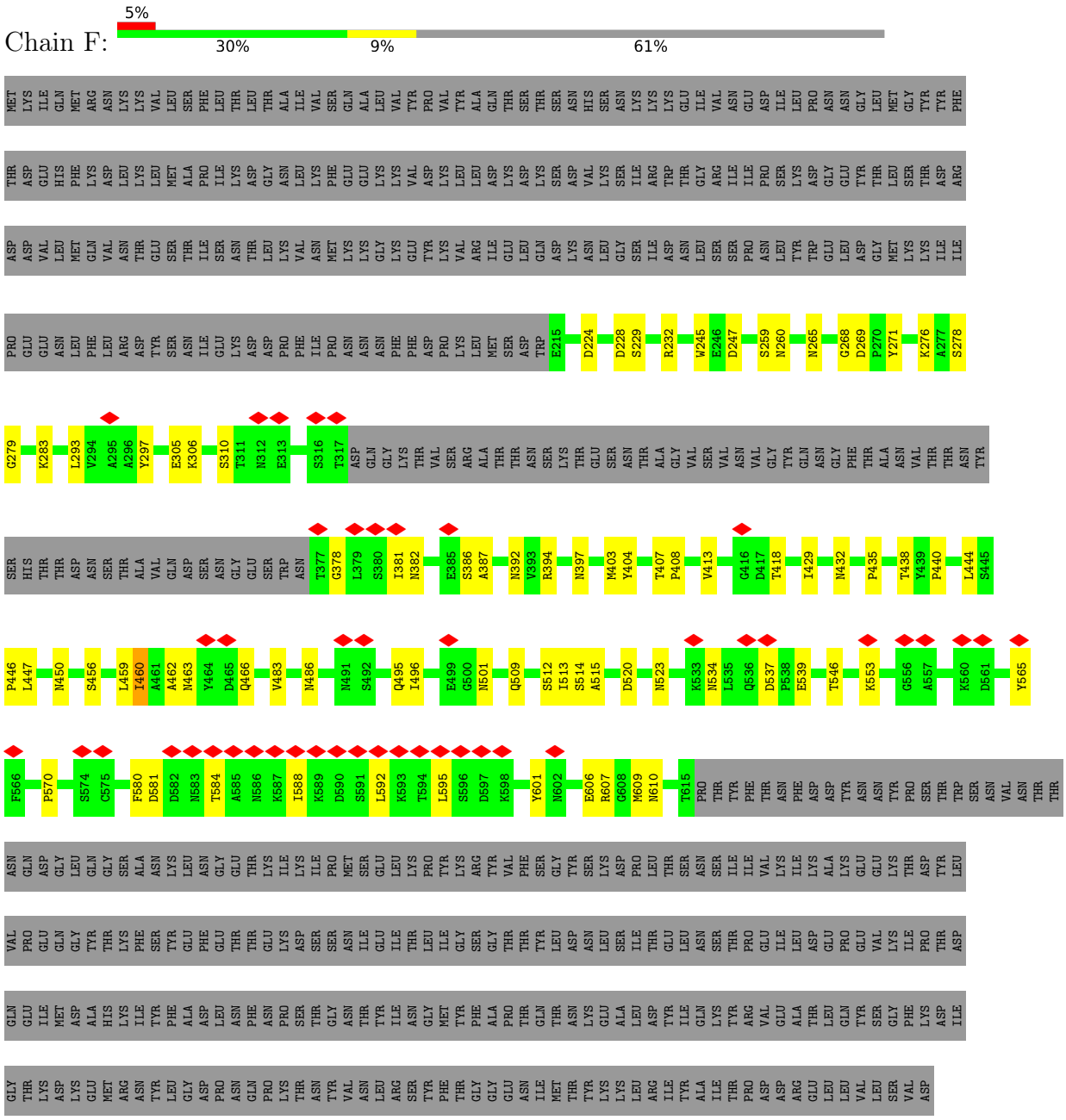


• Molecule 1: ADP-ribosyltransferase binding component

Chain E: 5% 30% 9% 61%



- Molecule 1: ADP-ribosyltransferase binding component



- Molecule 1: ADP-ribosyltransferase binding component

Chain G: 5% 30% 8% 61%

MET LYS ILE GLN MET ARG ASN LYS LYS VAL LEU SER PHE LEU THR LEU THR ALA ILE VAL SER GLN ALA LEU VAL TYR PRO VAL TYR ALA GLN THR SER THR SER ASN HIS SER ASN LYS LYS LYS GLU ILE VAL ASN GLU ASP ILE LEU PRO ASN ASN GLY LEU MET GLY TYR TYR PHE

THR ASP GLU HIS PHE LYS ASP LEU LYS LEU MET ALA PRO ILE LYS ASP GLY ASN LEU LYS PHE GLU GLU LYS LYS VAL ASP LYS LYS LEU LEU ASP LYS ASP LYS SER ASP VAL LYS SER ILE ARG TRP THR GLY ARG ILE ILE PRO SER LYS ASP GLY GLU TYR THR LEU SER THR ASP ARG

ASP ASP VAL LEU MET GLN VAL ASN THR GLU SER THR ILE SER ASN THR LEU LYS VAL ASN MET LYS LYS GLY LYS GLU TYR LYS VAL ARG ILE GLU LEU GLN ASP LYS ASN LEU GLY SER ILE ASP ASN LEU SER SER PRO ASN LEU TYR TRP GLU LEU ASP GLY MET LYS LYS ILE ILE

PRO GLU GLU ASN LEU PHE LEU ARG ASP TYR SER ASN ILE GLU LYS ASP ASP PRO PHE ILE PRO ASN ASN ASN PHE PHE ASP PRO LYS LEU MET SER ASP TRP **E215** **D224** **N225** **D228** **S229** **R232** **I241** **W245** **E246** **D247** **S259** **N260** **N265** **T266** **A267** **G268** **D269** **S278**

G279 L293 V294 A295 A296 Y297 E305 K306 S310 T311 N312 E313 S316 T317 ASP GLN GLY LYS THR VAL SER ARG ALA THR THR ASN SER LYS THR GLU SER ASN THR ALA GLY VAL SER VAL ASN VAL GLY TYR GLN ASN GLY PHE THR ALA ASN VAL THR THR ASN TYR SER HIS

THR THR ASP ASN SER THR ALA VAL GLN ASP SER ASN GLY GLU SER TRP ASN T377 G378 L379 S380 I381 E385 S386 A387 N392 V393 R394 N397 Y404 T407 P408 V413 G416 D417 T418 I429 G430 N431 N432 P435 T438 Y439 P440 L444 S445 P446 N450

S456 S457 R458 L459 I460 A461 A462 N463 Y464 D465 Q466 V483 F487 N491 S492 Q495 I496 E499 W503 S512 I513 S514 A515 D520 T521 E522 K533 Q536 D537 P538 E539 E552 K553 G556 K560 D561 Y565 F566 P570 I571 S574 C575

F580 D581 D582 N583 T584 A585 N586 K587 I588 K589 D590 S591 L592 K593 T594 L595 S596 D597 K598 N602 E606 R607 G608 M609 N610 T615 PRO THR TYR PHE THR ASN PHE ASP ASP TYR ASN ASN TYR PRO SER THR TRP SER ASN VAL ASN THR THR ASN GLN ASP GLY LEU GLN GLY SER

ALA ASN LYS LEU ASN GLY GLU THR LYS ILE LYS ILE PRO MET SER GLU LEU LYS PRO TYR LYS ARG TYR VAL PHE SER GLY TYR SER LYS ASP PRO LEU THR SER ASN SER ILE ILE VAL LYS ILE LYS ALA LYS GLU GLU LYS THR ASP TYR LEU VAL PRO GLU GLN GLY TYR THR LYS

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

ILE TYR PHE ALA ASP LEU ASN PHE ASN PRO SER THR GLY ASN THR TYR ILE ASN GLY MET TYR PHE ALA PRO THR GLN THR ASN LYS GLU ALA LEU ASP TYR ILE GLN LYS TYR ARG VAL GLU ALA THR LEU GLN TYR SER GLY PHE LYS ASP ILE GLY THR LYS ASP LYS GLU MET ARG

ASN TYR LEU GLY ASP PRO ASN GLN PRO LYS THR ASN TYR VAL ASN LEU ARG SER TYR PHE THR GLY GLY GLU ASN ILE MET THR TYR LYS LYS LEU ARG ILE TYR ALA ILE THR PRO ASP ASP ARG GLU LEU LEU VAL LEU SER VAL ASP

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C7	Depositor
Number of particles used	19844	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 ($e^-/Å^2$)	73.44	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	0.048	Depositor
Minimum map value	-0.038	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.01	Depositor
Map size (Å)	248.6, 248.6, 248.6	wwPDB
Map dimensions	220, 220, 220	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.13, 1.13, 1.13	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	0.38	0/2699	0.61	0/3655
1	B	0.38	0/2699	0.61	0/3655
1	C	0.38	0/2699	0.61	0/3655
1	D	0.38	0/2699	0.61	0/3655
1	E	0.38	0/2699	0.62	0/3655
1	F	0.38	0/2699	0.61	0/3655
1	G	0.38	0/2699	0.61	0/3655
All	All	0.38	0/18893	0.61	0/25585

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	2
1	B	0	2
1	C	0	2
1	D	0	2
1	E	0	2
1	F	0	2
1	G	0	2
All	All	0	14

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	456	SER	Peptide
1	A	460	ILE	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	456	SER	Peptide
1	B	460	ILE	Peptide
1	C	456	SER	Peptide

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	2659	0	2602	50	0
1	B	2659	0	2602	54	0
1	C	2659	0	2602	50	0
1	D	2659	0	2602	52	0
1	E	2659	0	2602	55	0
1	F	2659	0	2602	54	0
1	G	2659	0	2602	51	0
All	All	18613	0	18214	348	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 9.

The worst 5 of 348 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:460:ILE:HG22	1:E:466:GLN:HB3	1.79	0.64
1:A:460:ILE:HG22	1:A:466:GLN:HB3	1.81	0.62
1:D:460:ILE:HG22	1:D:466:GLN:HB3	1.82	0.61
1:G:460:ILE:HG22	1:G:466:GLN:HB3	1.84	0.60
1:F:460:ILE:HG22	1:F:466:GLN:HB3	1.84	0.59

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	338/876 (39%)	288 (85%)	50 (15%)	0	100	100
1	B	338/876 (39%)	286 (85%)	52 (15%)	0	100	100
1	C	338/876 (39%)	286 (85%)	52 (15%)	0	100	100
1	D	338/876 (39%)	285 (84%)	53 (16%)	0	100	100
1	E	338/876 (39%)	287 (85%)	51 (15%)	0	100	100
1	F	338/876 (39%)	286 (85%)	52 (15%)	0	100	100
1	G	338/876 (39%)	285 (84%)	53 (16%)	0	100	100
All	All	2366/6132 (39%)	2003 (85%)	363 (15%)	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	297/788 (38%)	295 (99%)	2 (1%)	84	90
1	B	297/788 (38%)	294 (99%)	3 (1%)	76	86
1	C	297/788 (38%)	296 (100%)	1 (0%)	92	95
1	D	297/788 (38%)	293 (99%)	4 (1%)	69	82
1	E	297/788 (38%)	294 (99%)	3 (1%)	76	86
1	F	297/788 (38%)	294 (99%)	3 (1%)	76	86
1	G	297/788 (38%)	295 (99%)	2 (1%)	84	90

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2079/5516 (38%)	2061 (99%)	18 (1%)	79 87

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

Mol	Chain	Res	Type
1	F	382	ASN
1	G	580	PHE
1	G	297	TYR
1	D	497	VAL
1	F	297	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	475	GLN
1	F	486	ASN
1	E	486	ASN
1	F	397	ASN
1	G	397	ASN

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.

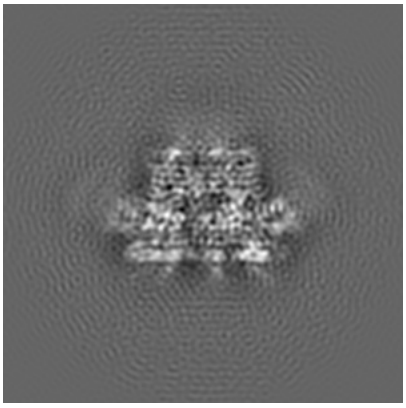
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-20102. These allow visual inspection of the internal detail of the map and identification of artifacts.

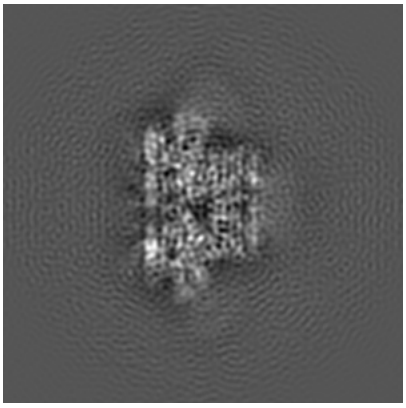
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

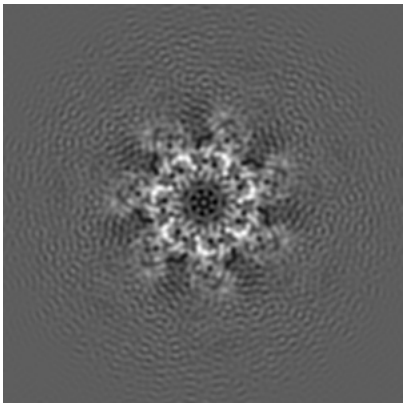
6.1.1 Primary map



X



Y

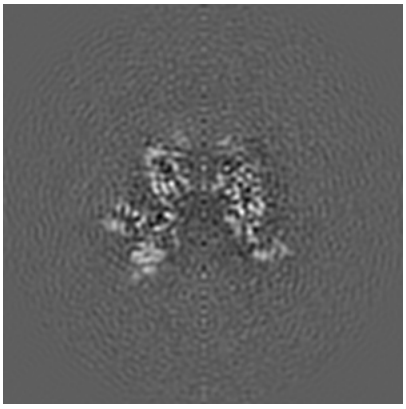


Z

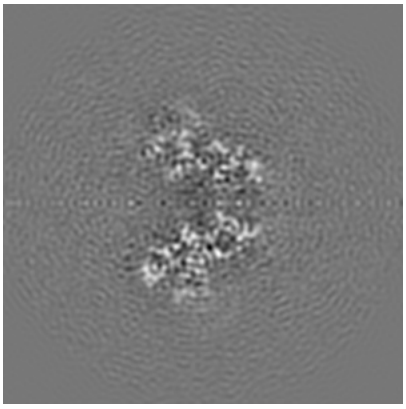
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

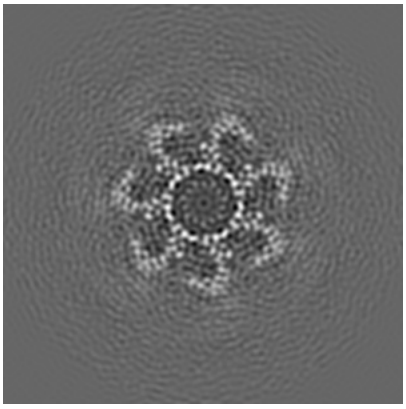
6.2.1 Primary map



X Index: 110



Y Index: 110

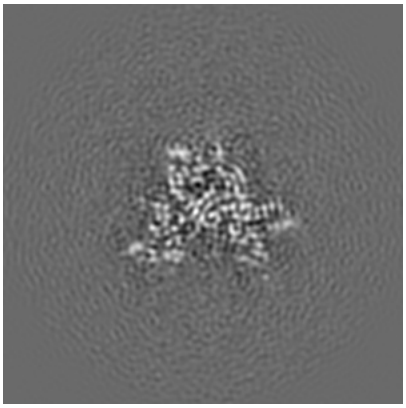


Z Index: 110

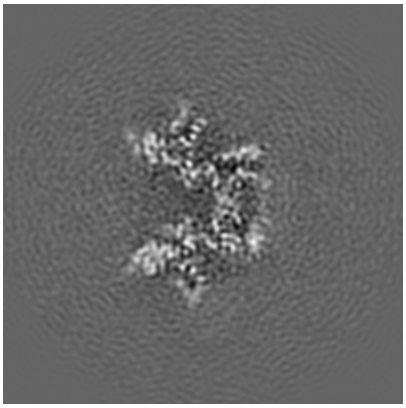
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

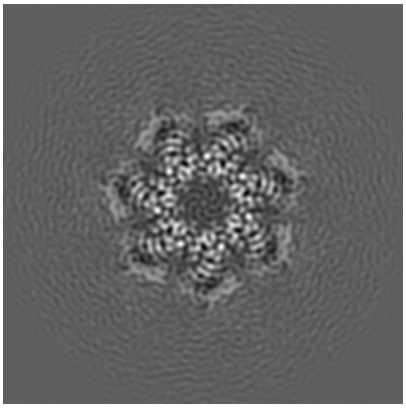
6.3.1 Primary map



X Index: 90



Y Index: 118

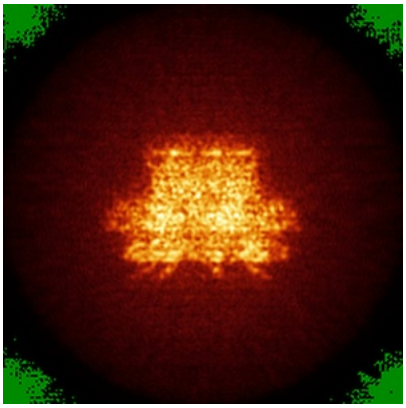


Z Index: 103

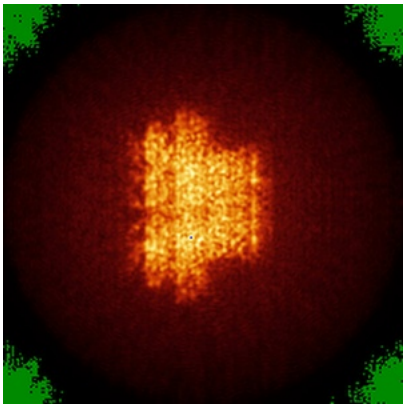
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) (i)

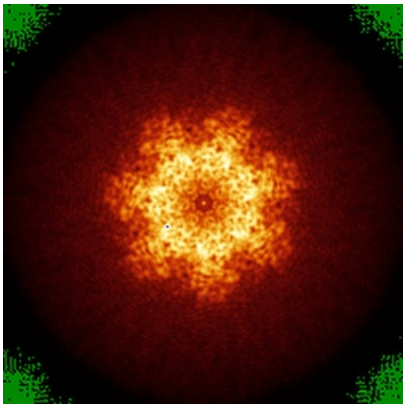
6.4.1 Primary map



X



Y

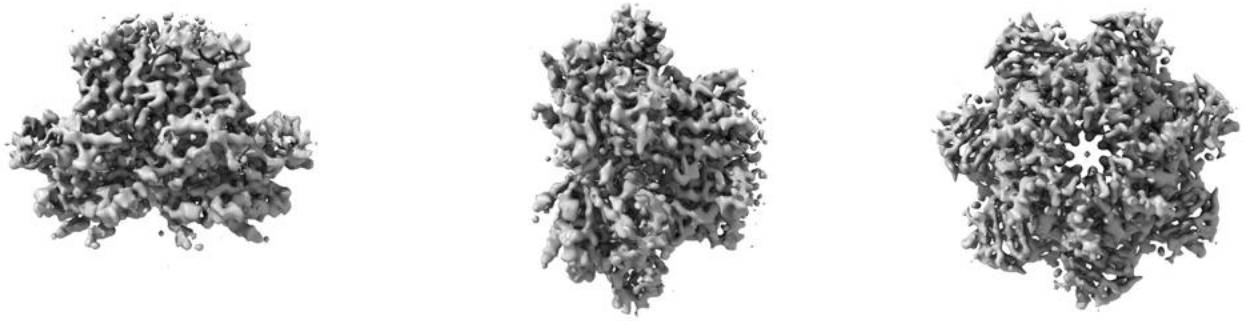


Z

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.01. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

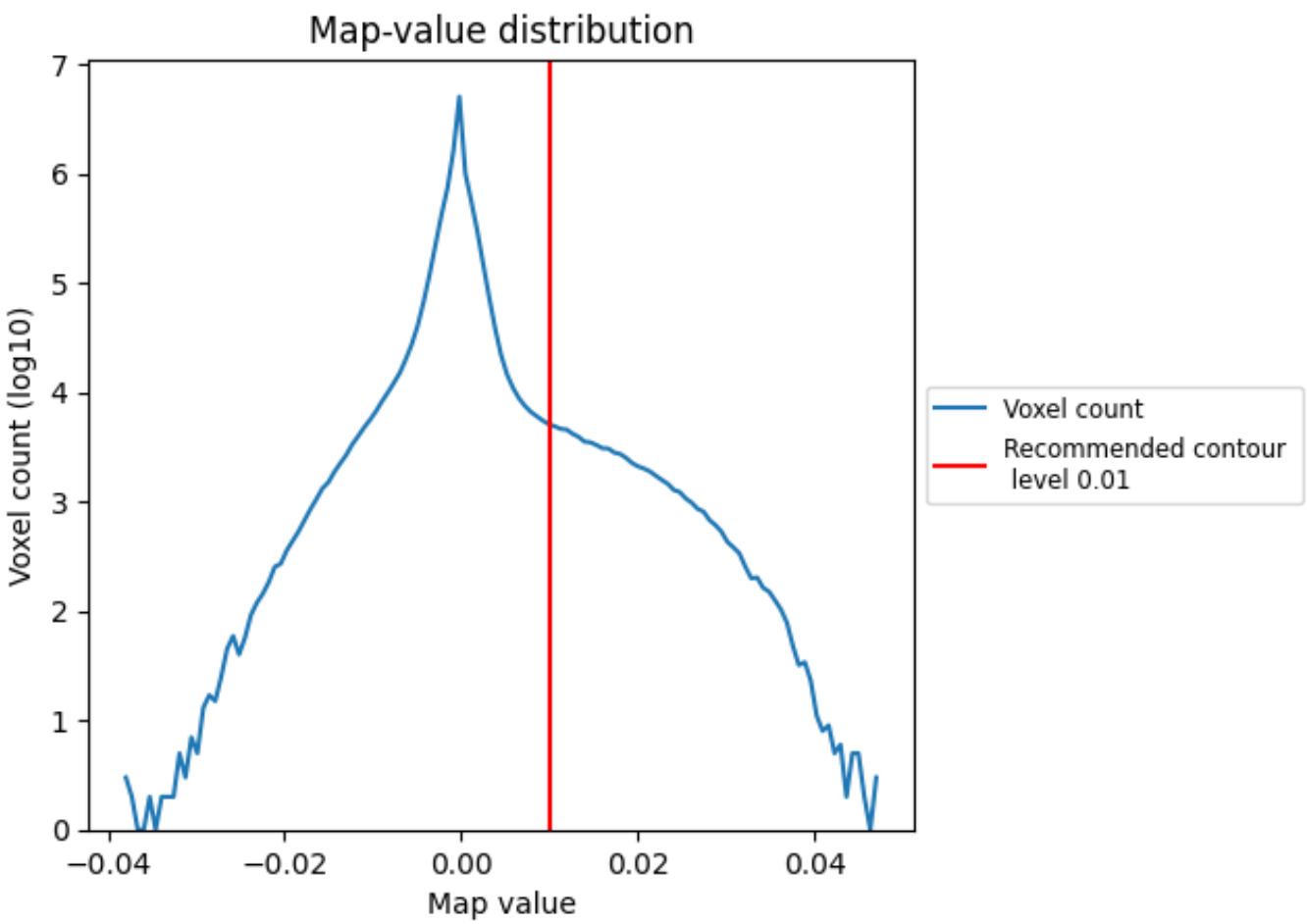
6.6 Mask visualisation (i)

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

7 Map analysis (i)

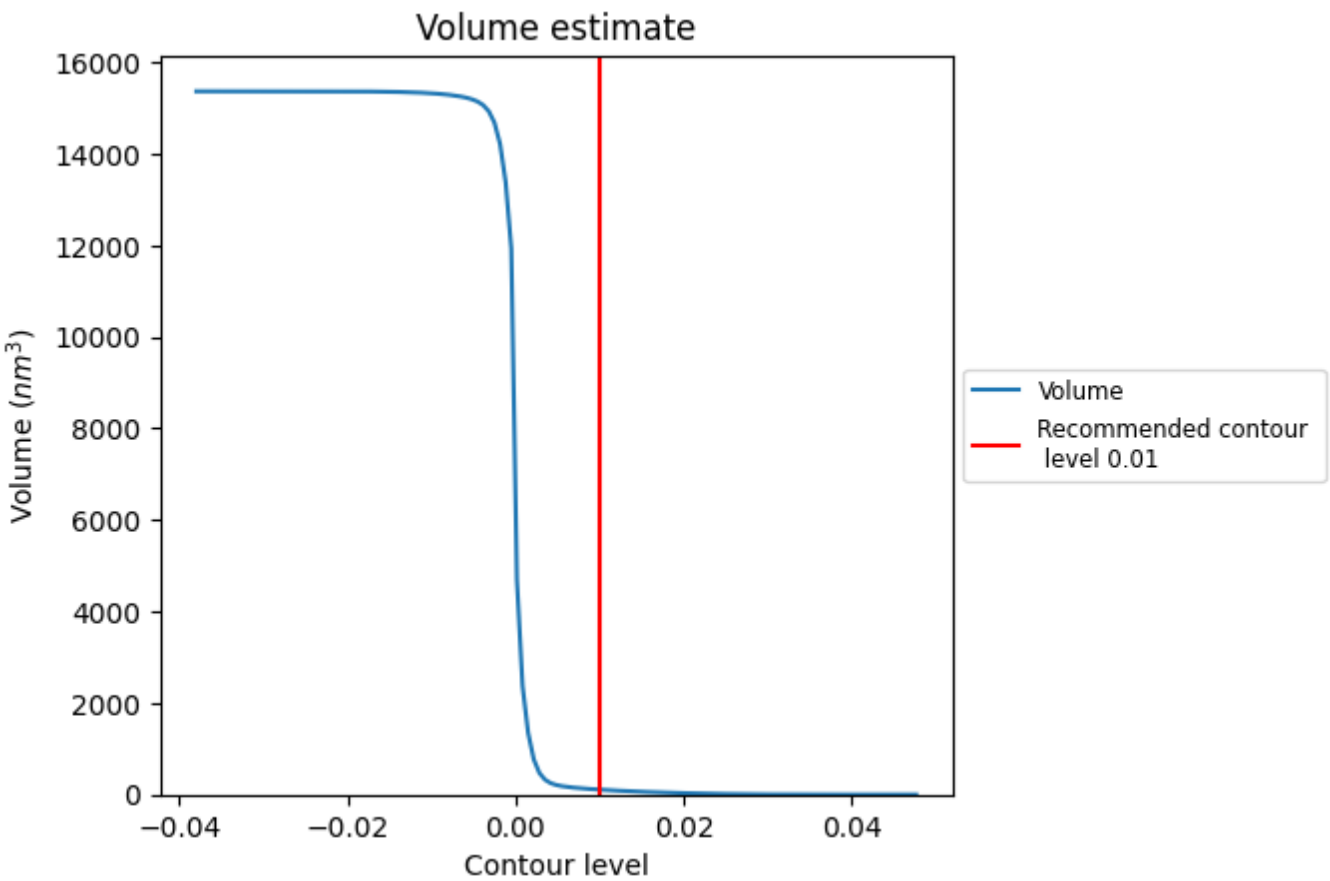
This section contains the results of statistical analysis of the map.

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.

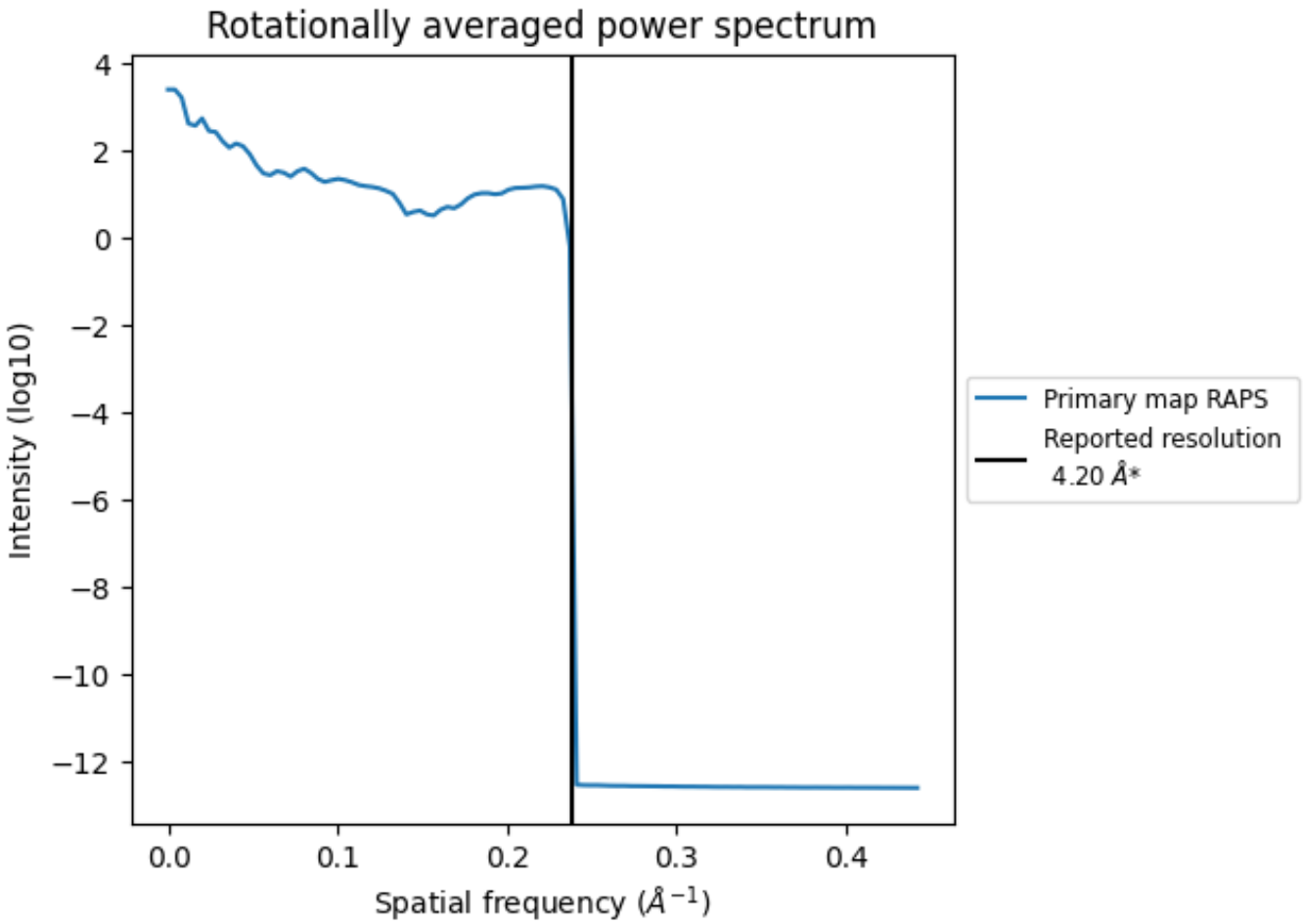
7.2 Volume estimate (i)



The volume at the recommended contour level is 108 nm^3; this corresponds to an approximate mass of 98 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)

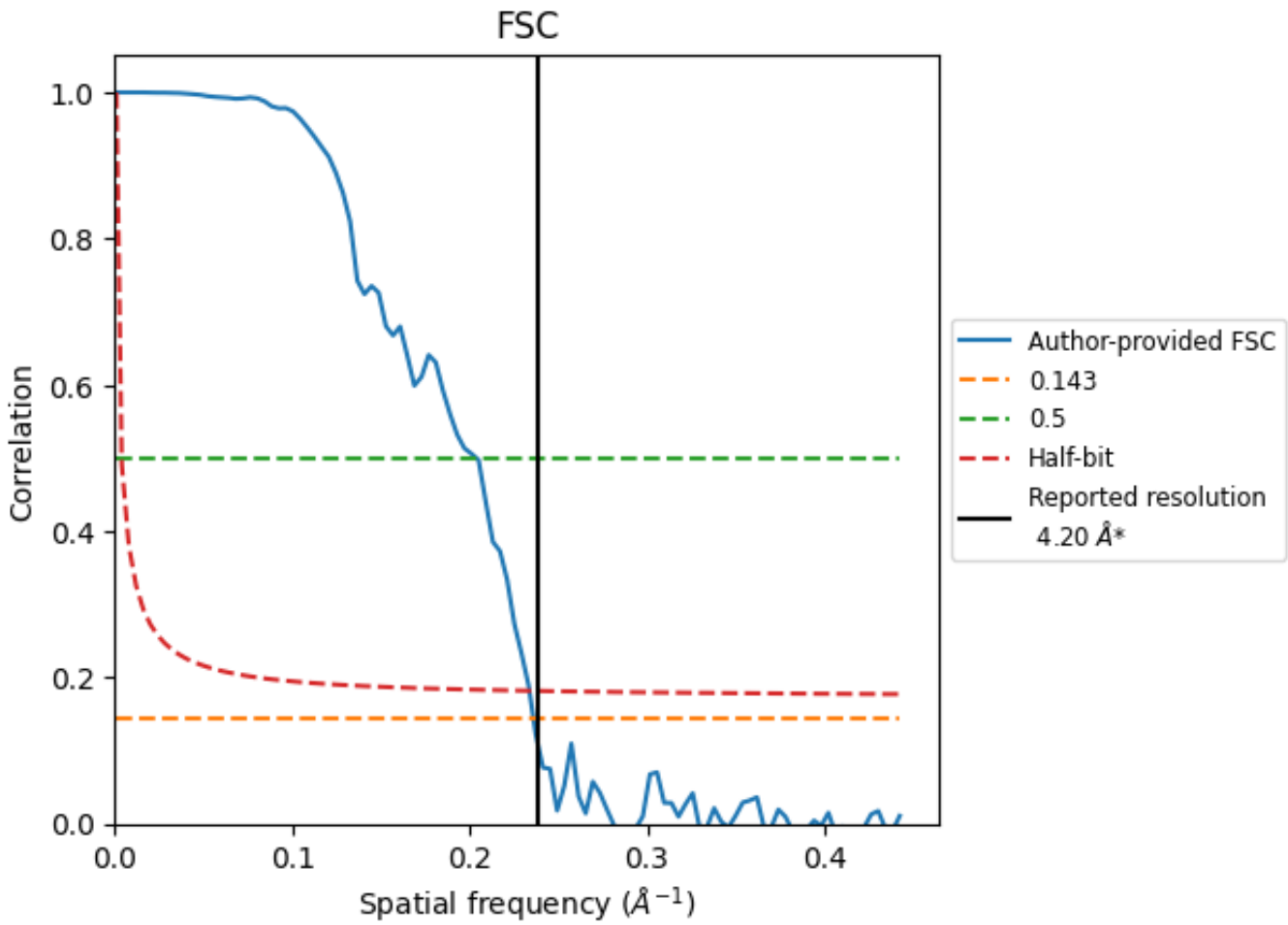


*Reported resolution corresponds to spatial frequency of 0.238 Å^{-1}

8 Fourier-Shell correlation (i)

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 (i)



*Reported resolution corresponds to spatial frequency of 0.238 Å^{-1}

8.2 Resolution estimates (i)

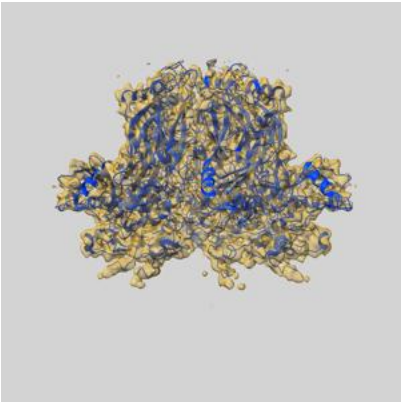
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.20	-	-
Author-provided FSC curve	4.24	4.91	4.28
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

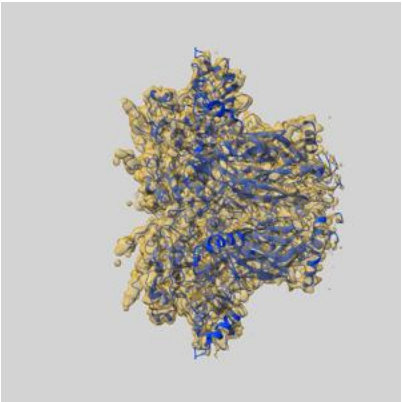
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-20102 and PDB model 6OKR. Per-residue inclusion information can be found in section 3 on page 4.

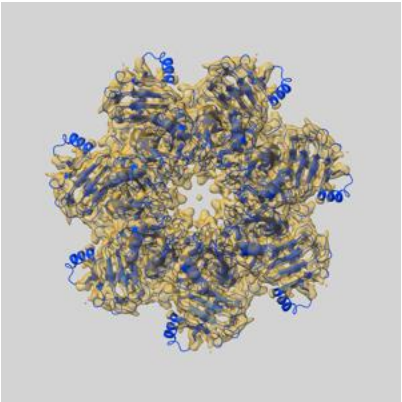
9.1 Map-model overlay (i)



X



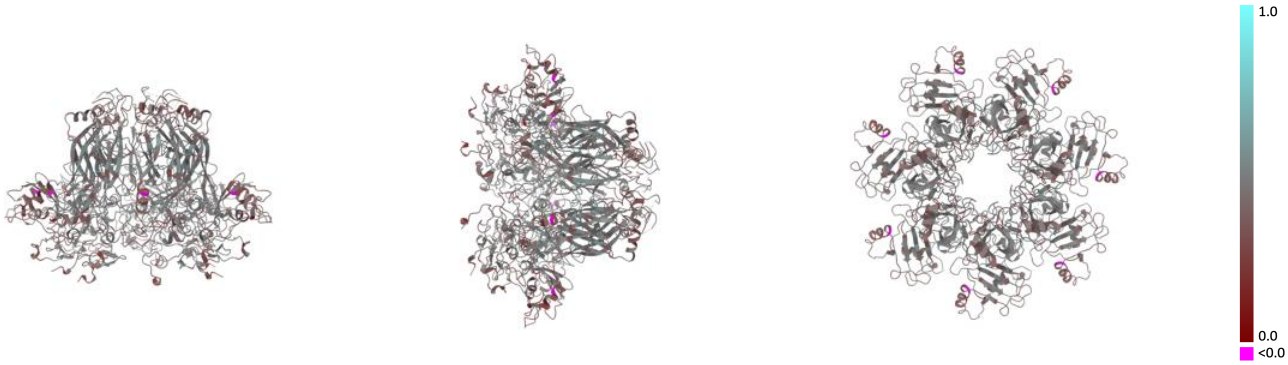
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.01 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.

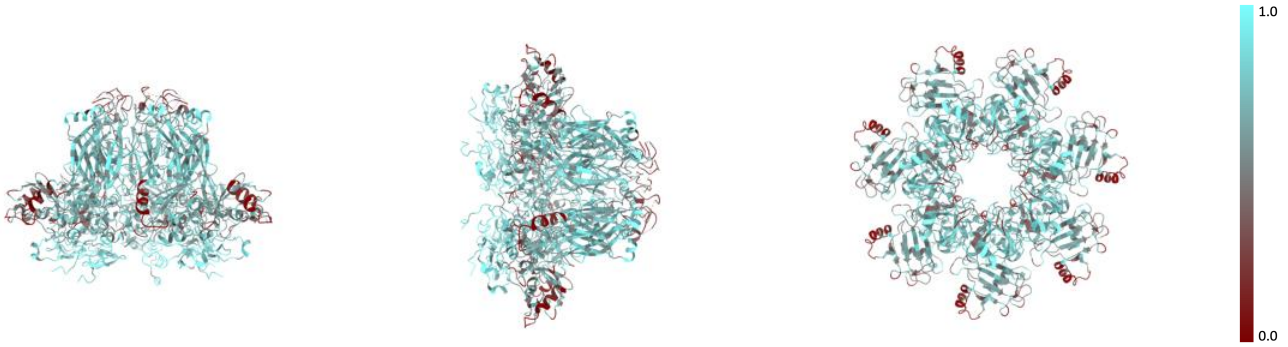
9.2 Q-score mapped to coordinate model (i)



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.

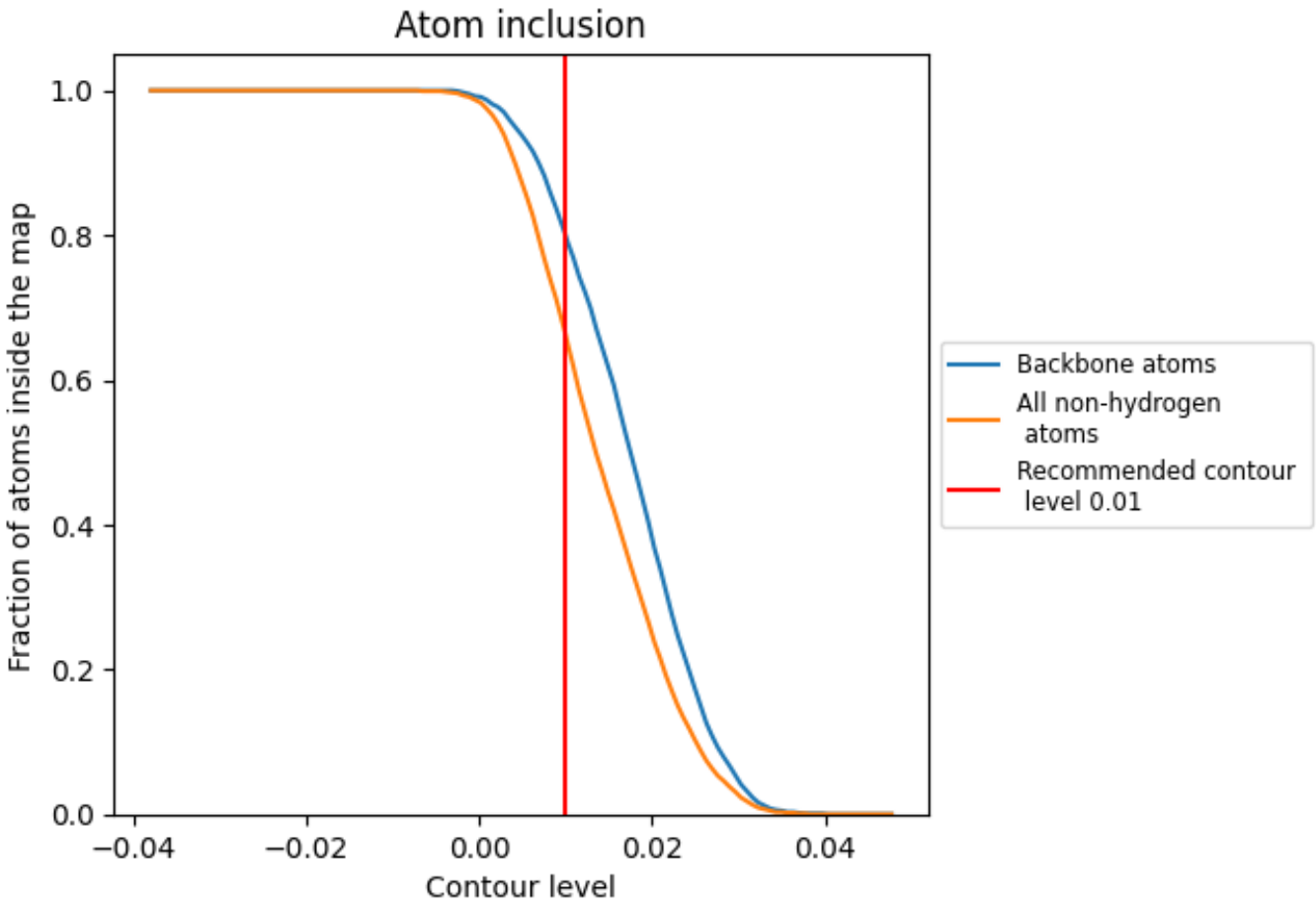
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.01).

9.4 Atom inclusion (i)



At the recommended contour level, 80% of all backbone atoms, 66% 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 (0.01) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6630	0.4340
A	0.6610	0.4340
B	0.6640	0.4330
C	0.6640	0.4330
D	0.6580	0.4340
E	0.6620	0.4350
F	0.6640	0.4360
G	0.6670	0.4350