wwPDB EM Validation Summary Report

Nov 8, 2022 – 08:23 AM EST

PDB ID	:	6OKS
EMDB ID	:	EMD-20103
Title	:	CDTb Double Heptamer Long Form Mask 1 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 ⓘ 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:

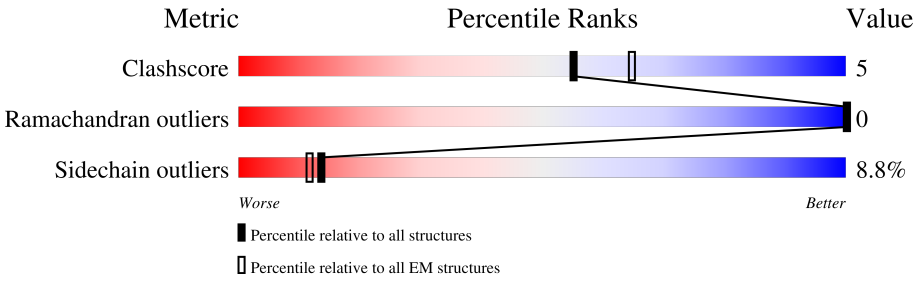
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (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	10% · 87%
1	B	876	10% · 87%
1	C	876	10% · 87%
1	D	876	11% · 87%
1	E	876	11% · 87%
1	F	876	11% · 87%
1	G	876	11% · 87%
1	H	876	11% · 87%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	876	
1	J	876	
1	K	876	
1	L	876	
1	M	876	
1	N	876	

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 13244 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.

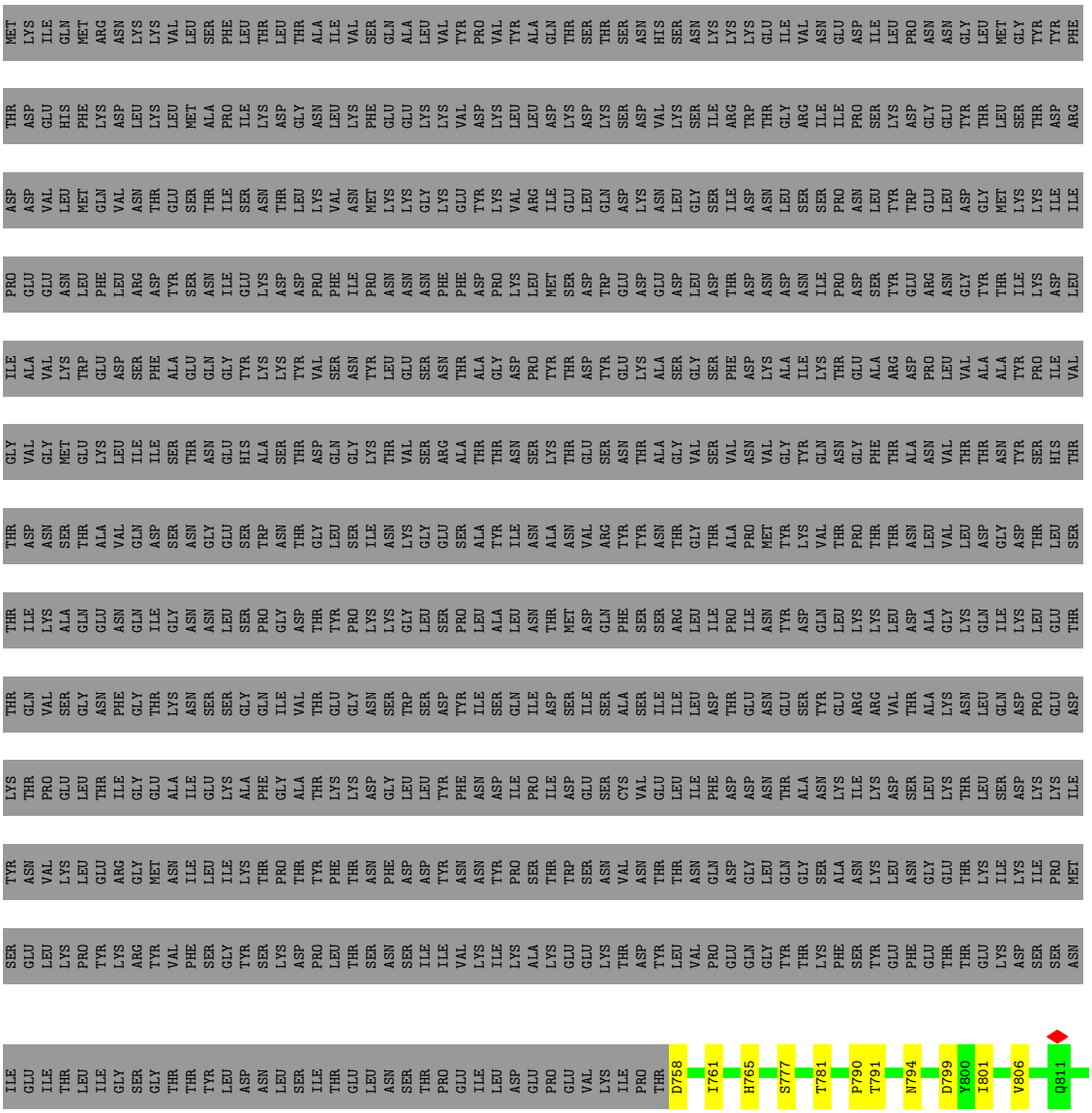
Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	B	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	C	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	D	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	E	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	F	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	G	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	H	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	I	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	J	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	K	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	L	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	M	115	Total	C	N	O	S	0	0
			946	602	156	184	4		
1	N	115	Total	C	N	O	S	0	0
			946	602	156	184	4		

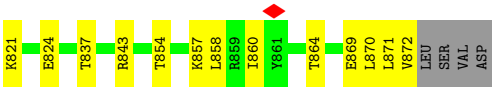
3 Residue-property plots (i)

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

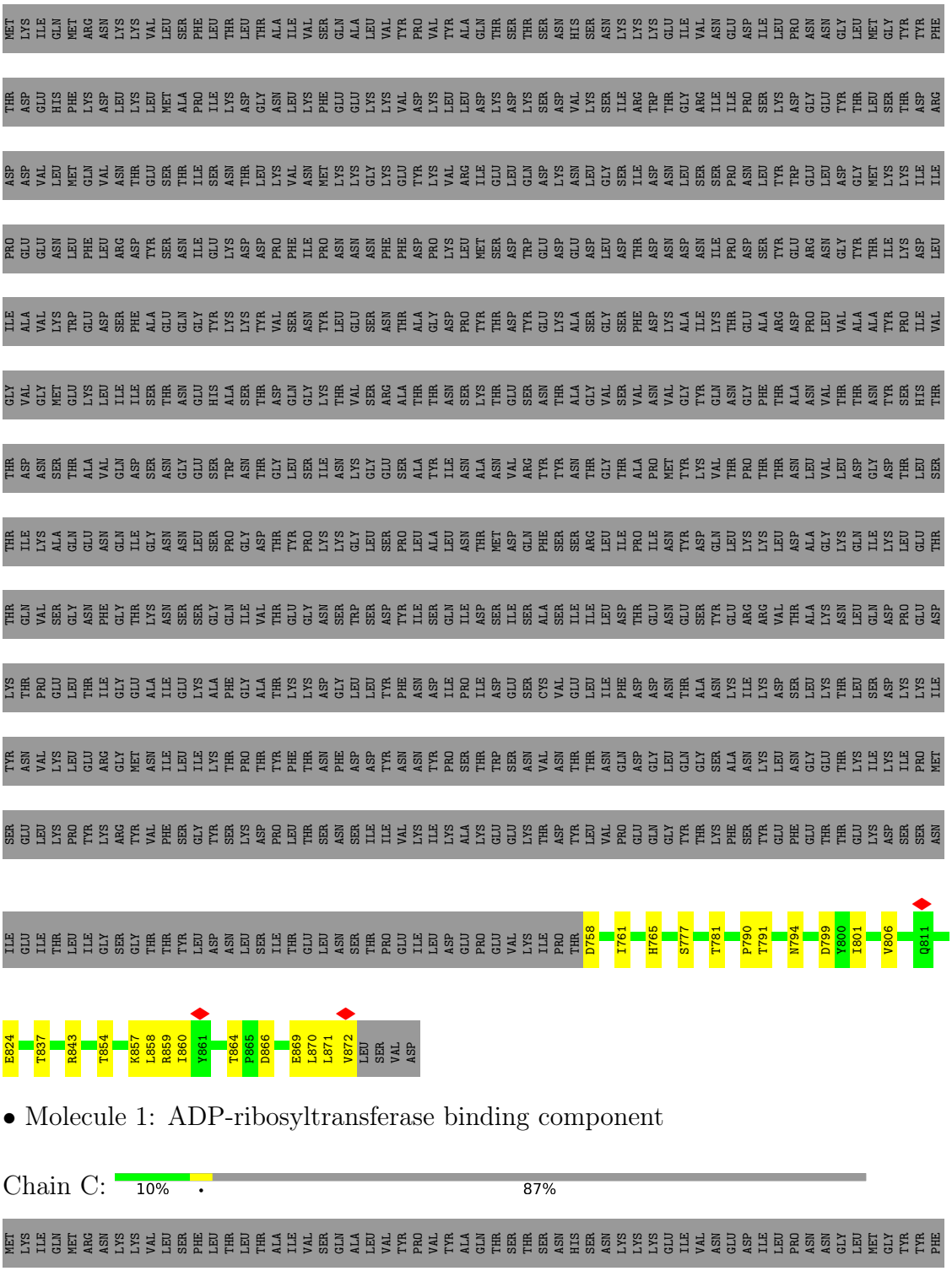
Chain A:





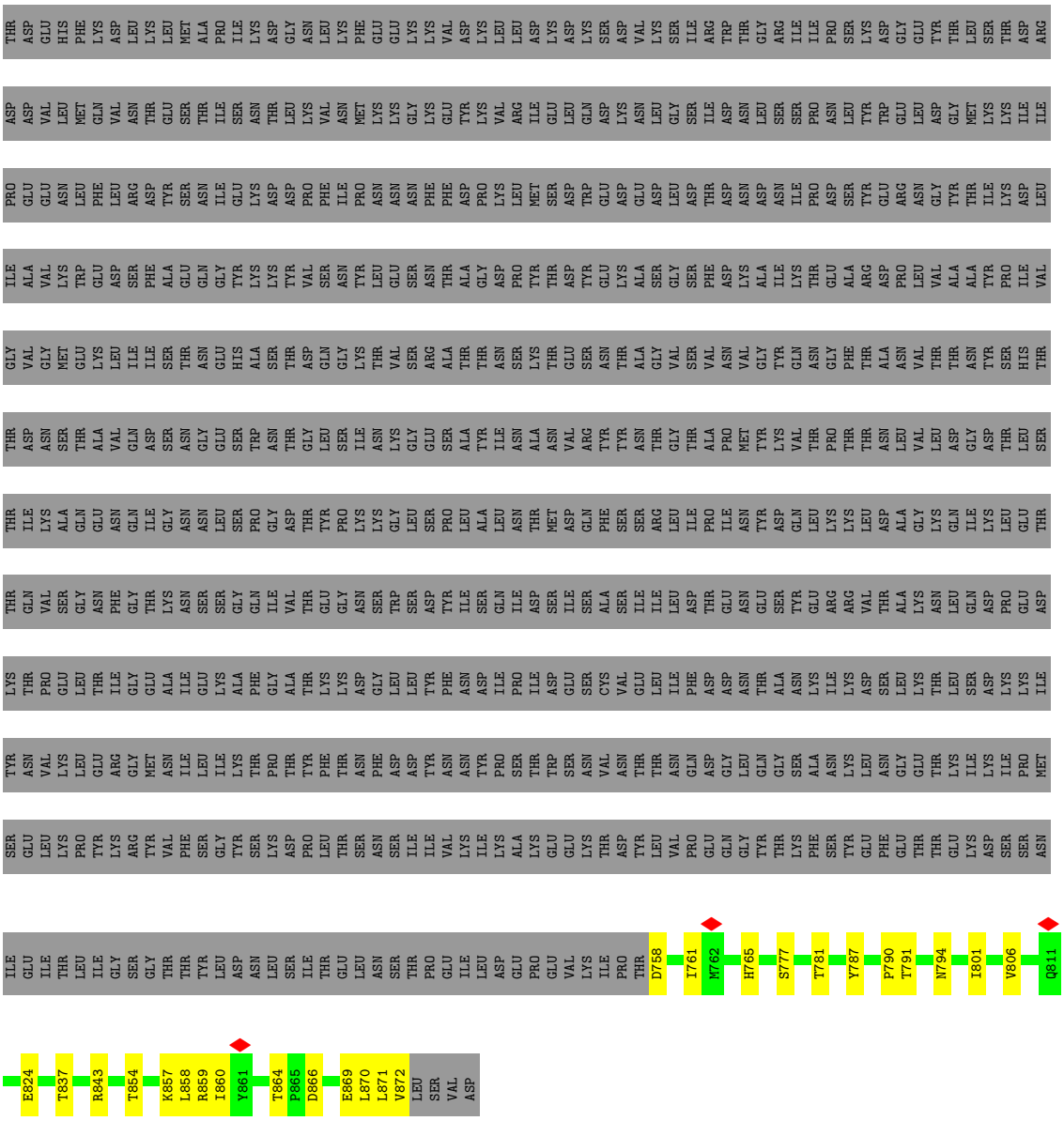
• Molecule 1: ADP-ribosyltransferase binding component

Chain B:



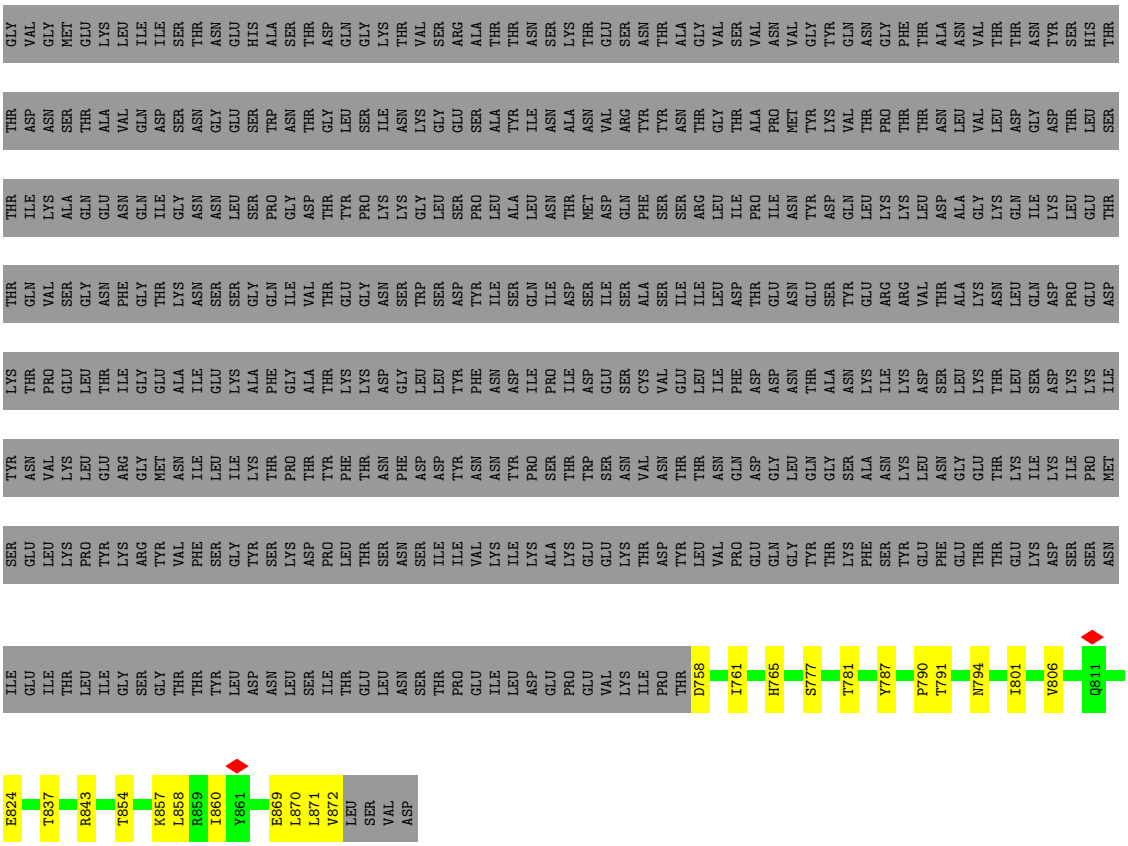
• Molecule 1: ADP-ribosyltransferase binding component

Chain C:



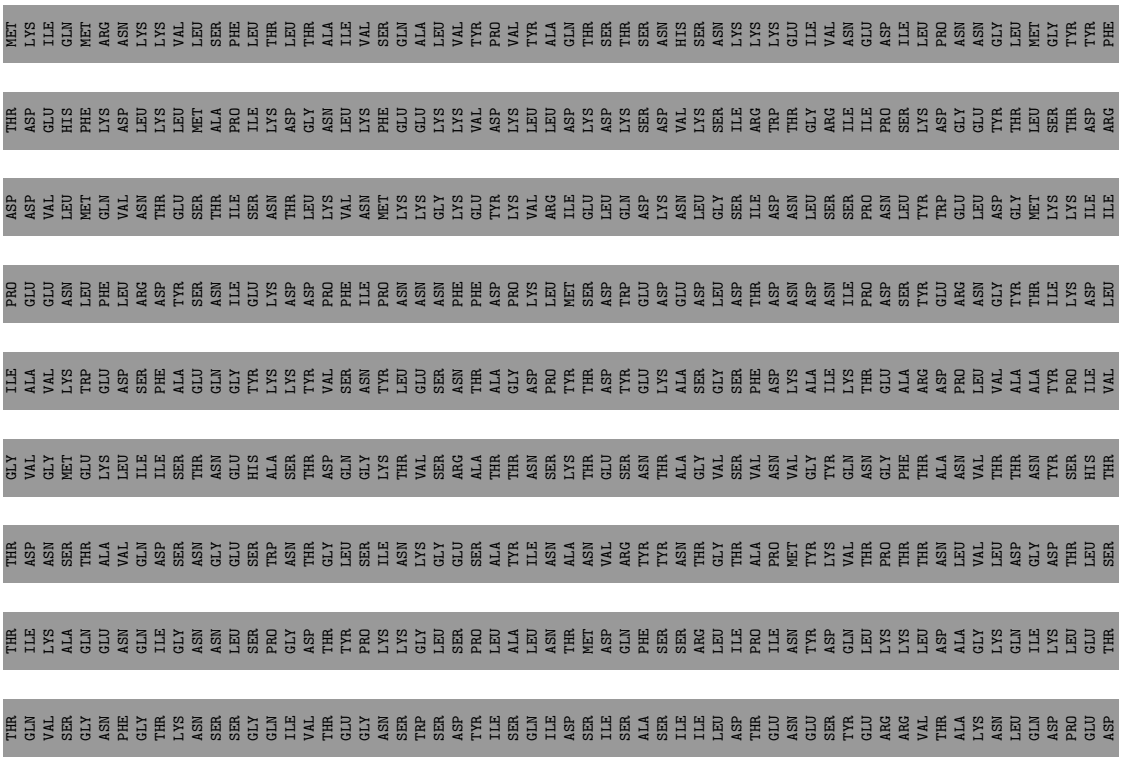
• Molecule 1: ADP-ribosyltransferase binding component

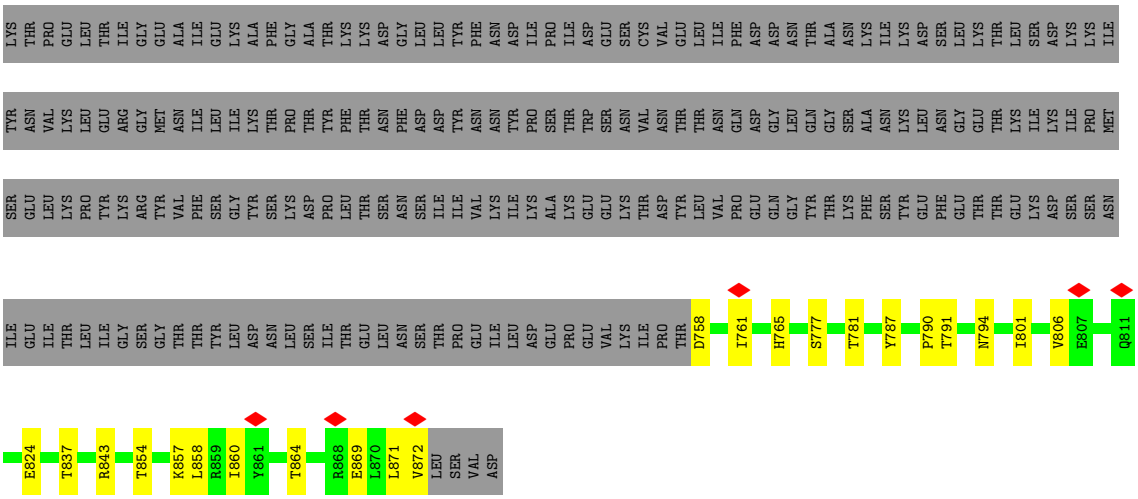
Chain D:



• Molecule 1: ADP-ribosyltransferase binding component

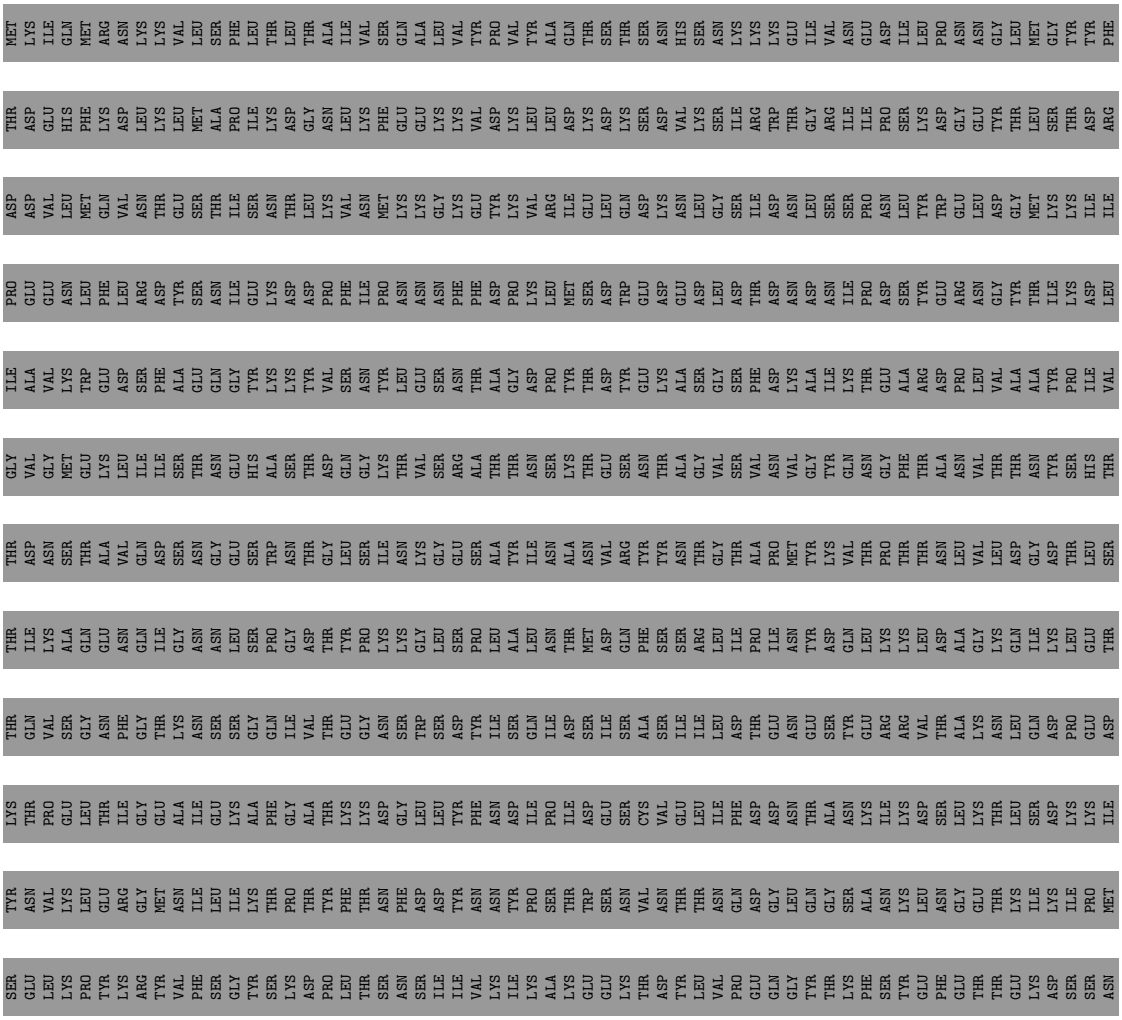
Chain E:

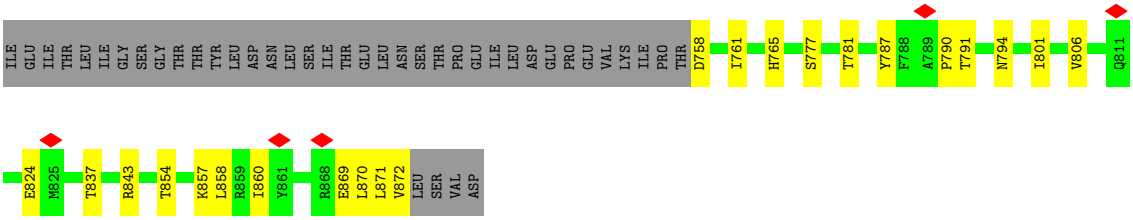




• Molecule 1: ADP-ribosyltransferase binding component

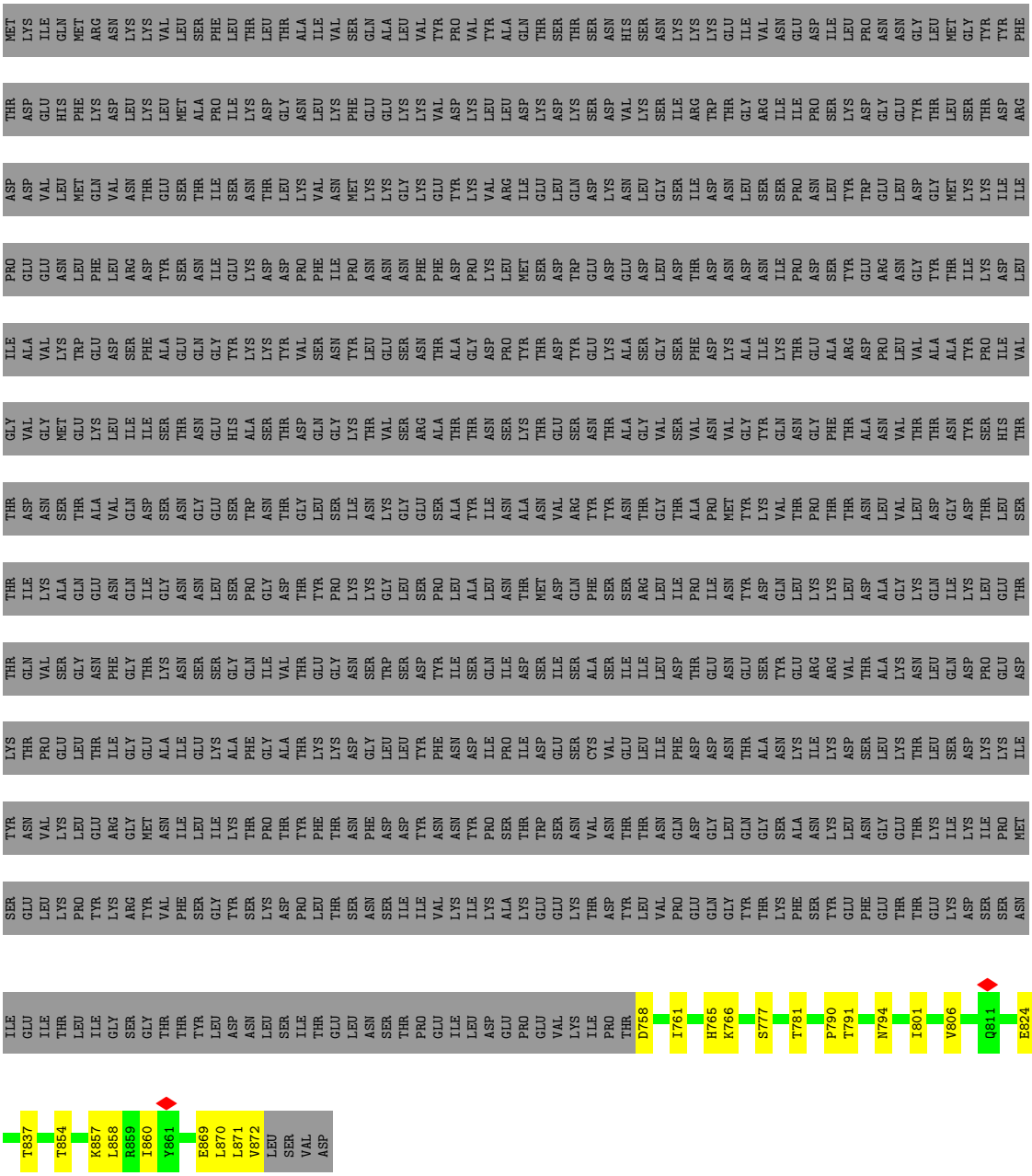
Chain F: 11% 87%





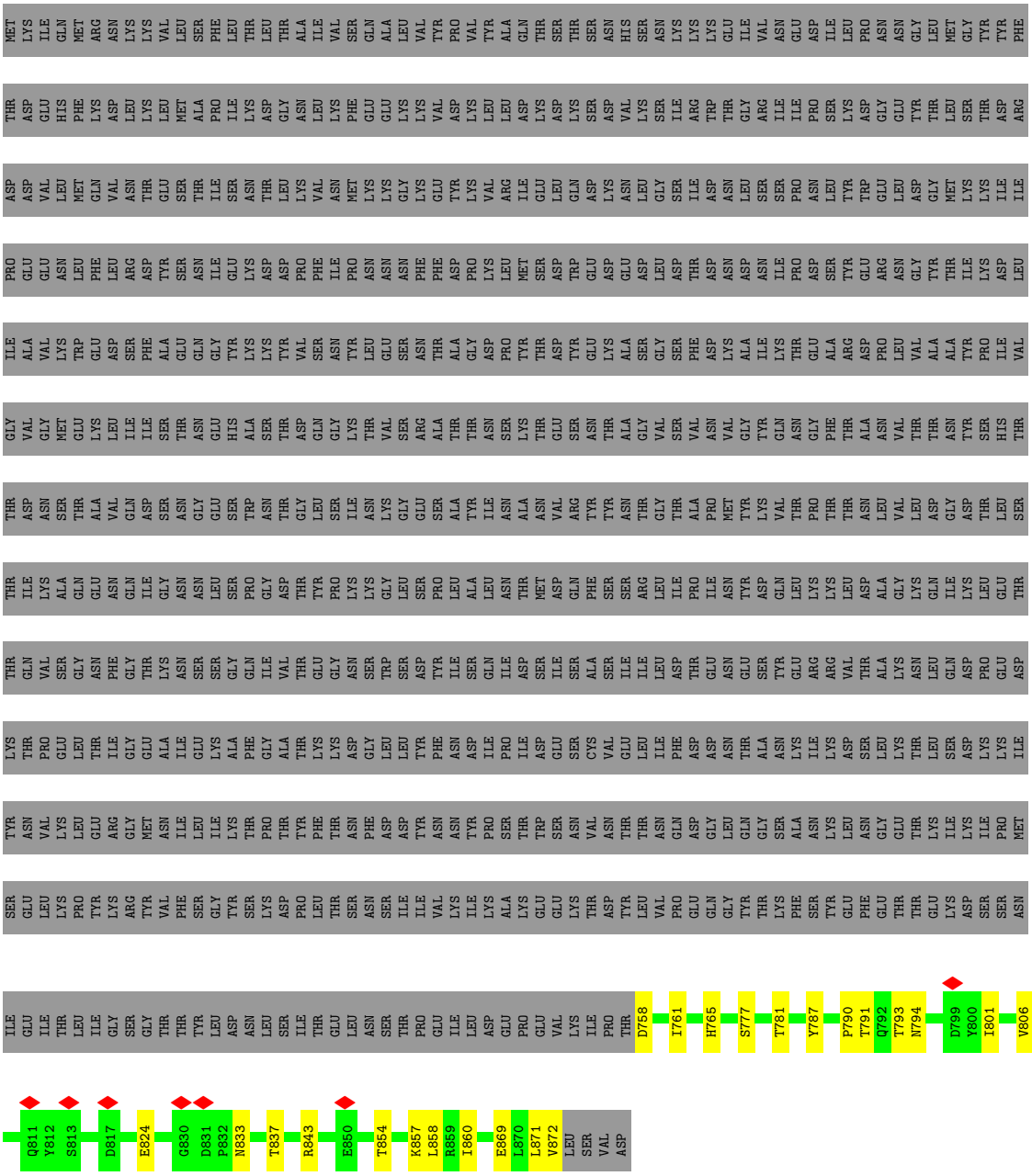
• Molecule 1: ADP-ribosyltransferase binding component

Chain G: 11% 87%



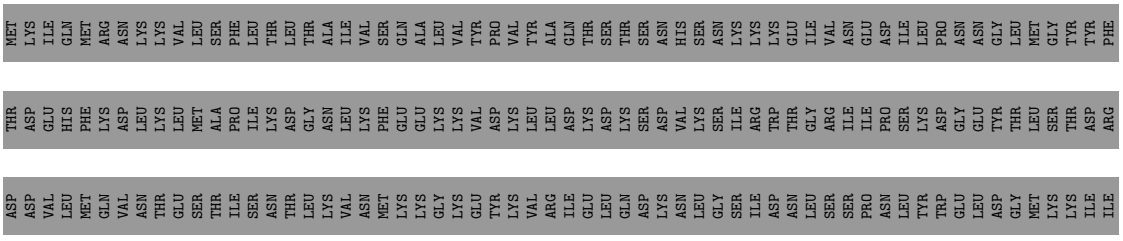
• Molecule 1: ADP-ribosyltransferase binding component

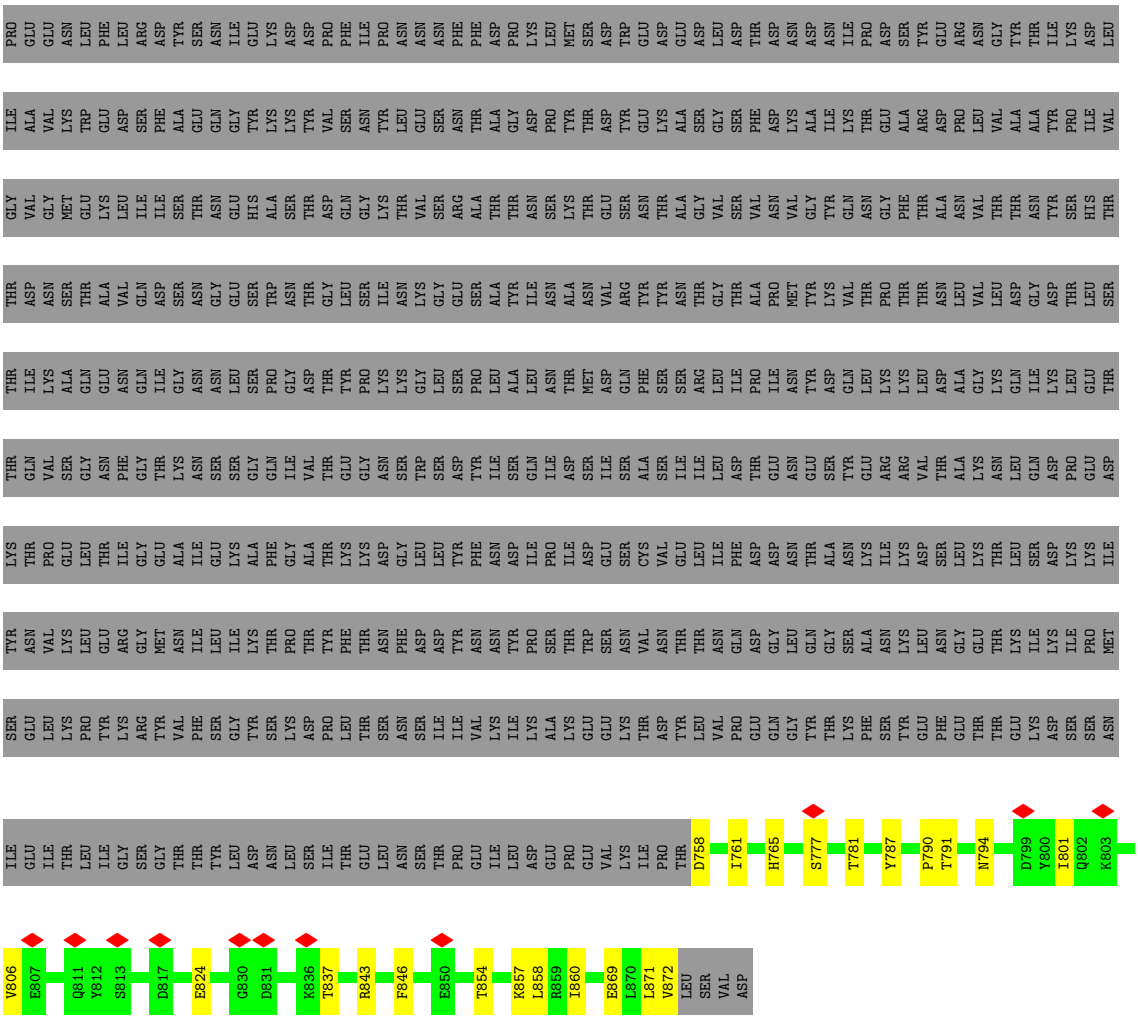
Chain H:



• Molecule 1: ADP-ribosyltransferase binding component

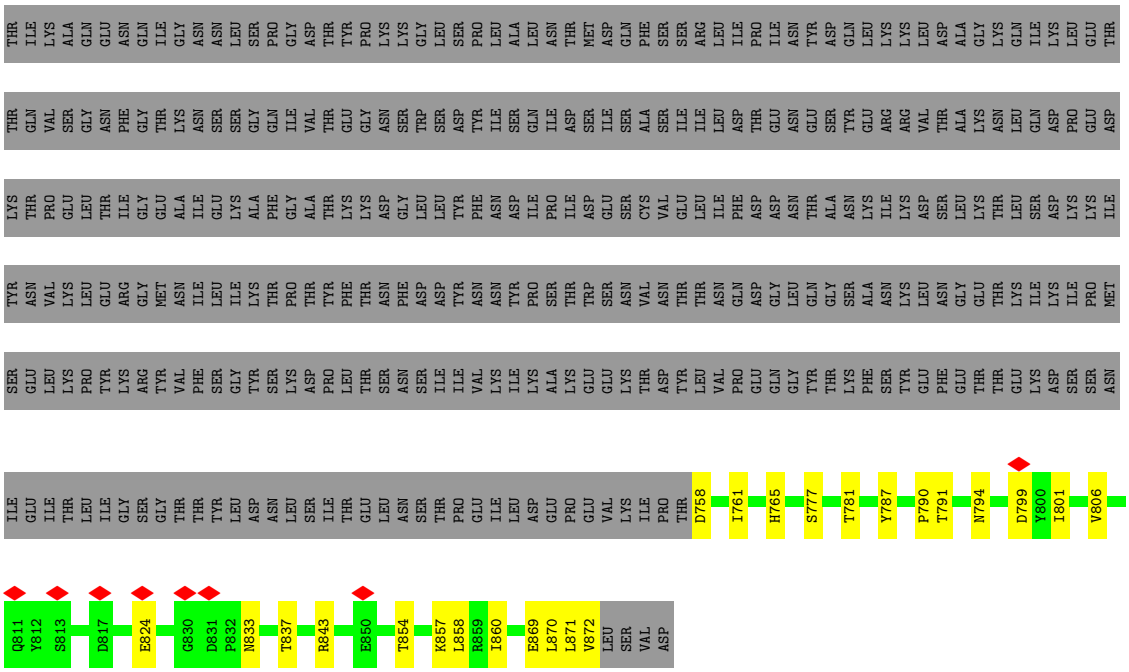
Chain I:



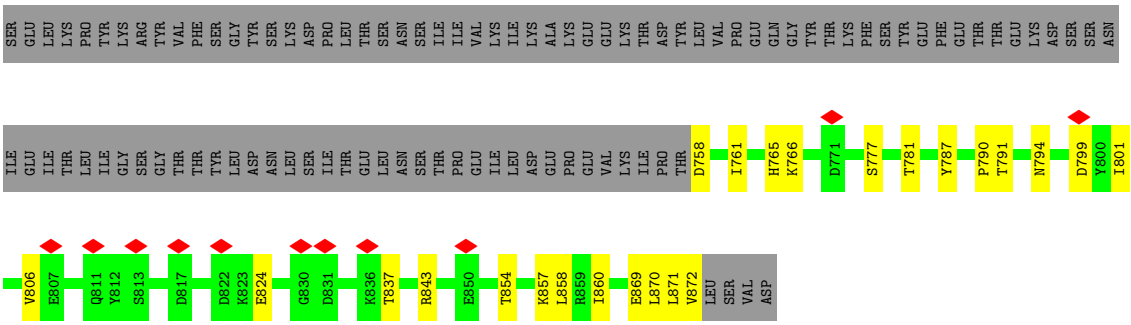


• Molecule 1: ADP-ribosyltransferase binding component

Chain J:

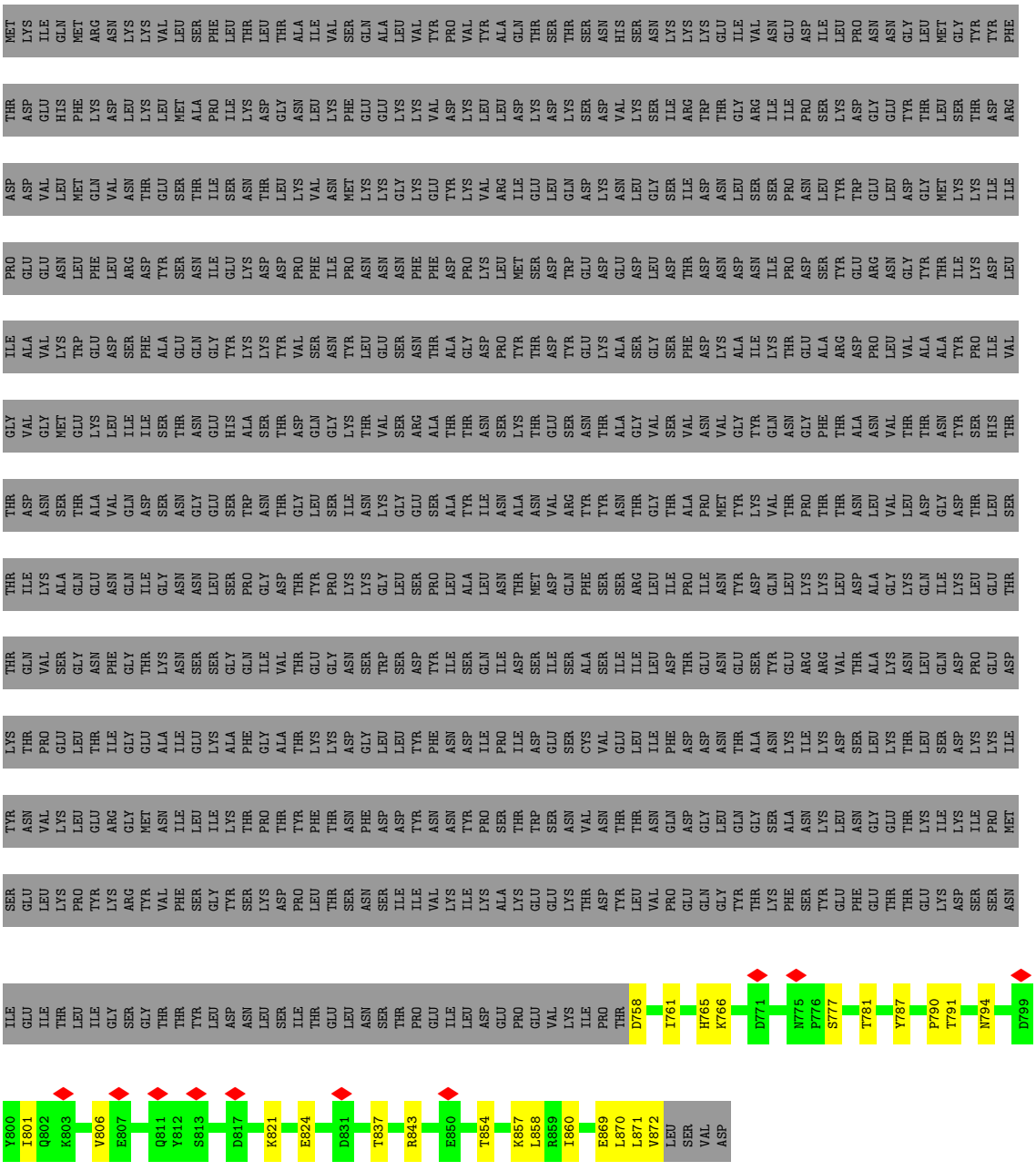


• Molecule 1: ADP-ribosyltransferase binding component



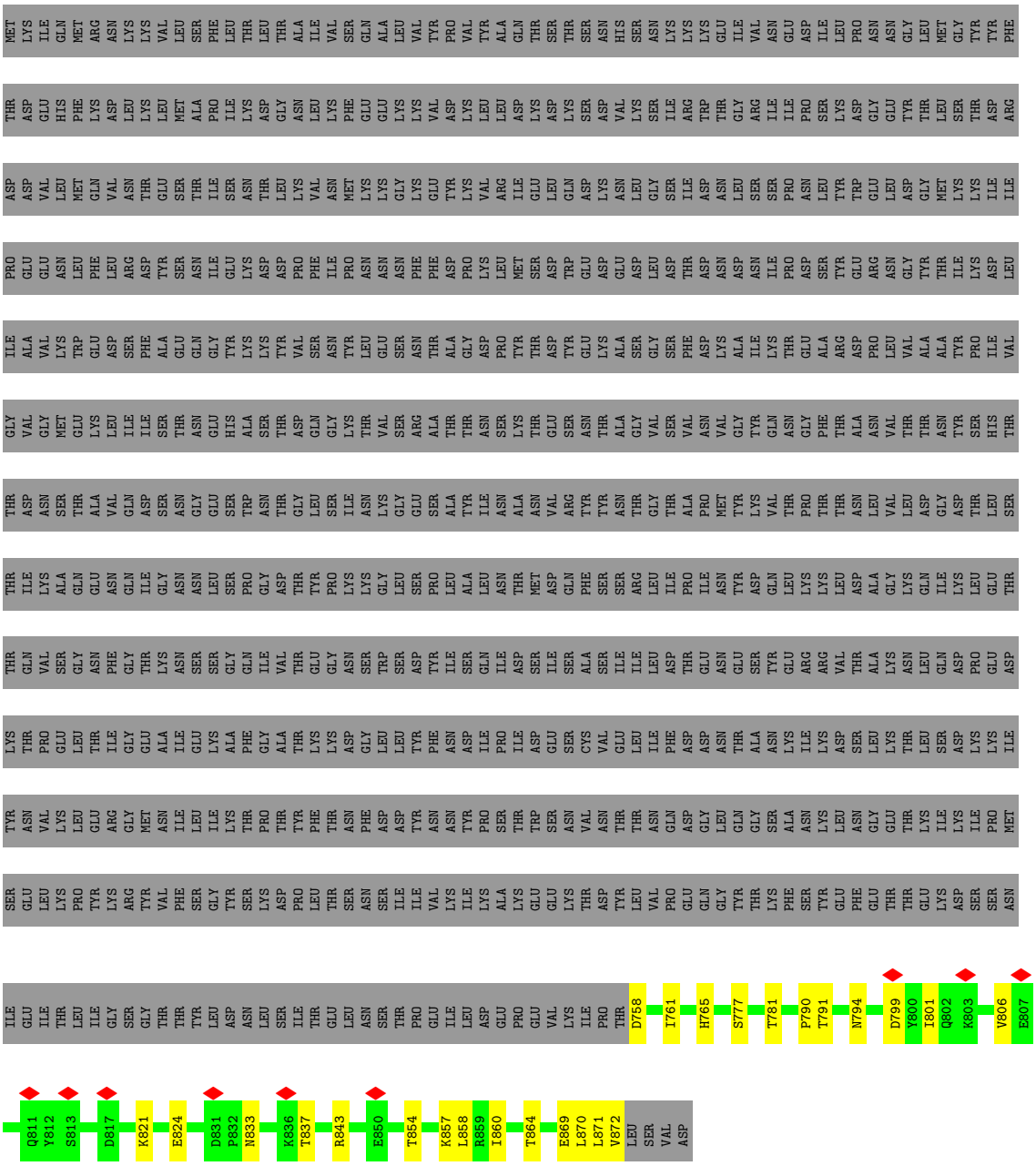
• Molecule 1: ADP-ribosyltransferase binding component

Chain L: 10% 87%



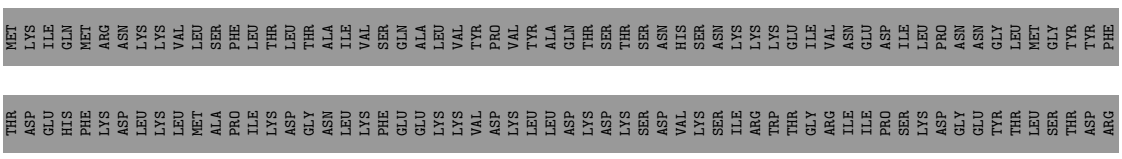
• Molecule 1: ADP-ribosyltransferase binding component

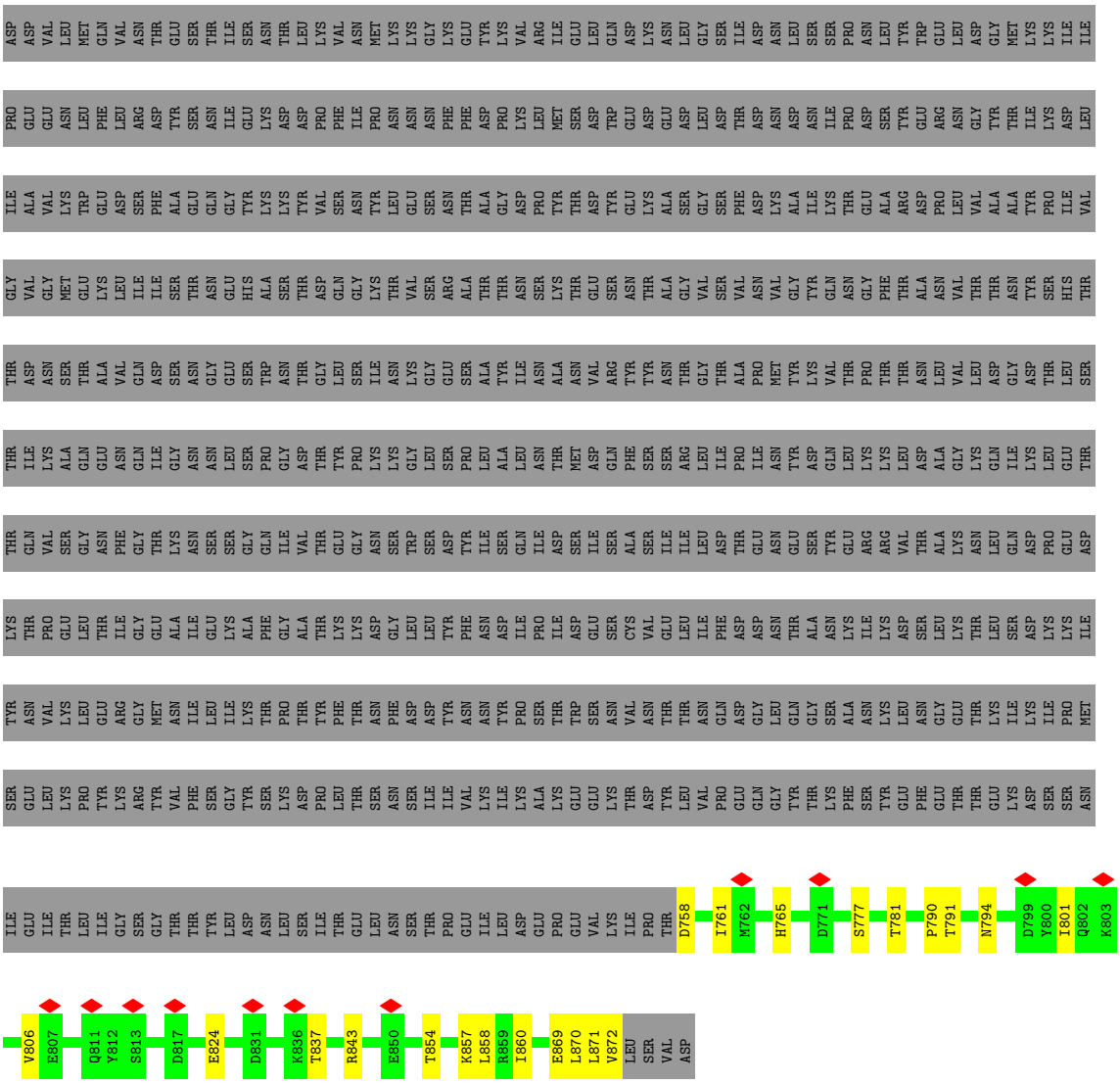
Chain M:



• Molecule 1: ADP-ribosyltransferase binding component

Chain N:




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 GLU LEU GLN ASP LYS ASN LEU GLY SER ILE ASP ASN LEU SER SER PRO ASN LEU TYR TRP GLU LEU SER 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 GLU ASP GLU ASP LEU ASP THR ASP ASN ASP ASN ILE PRO ASP SER TYR GLU ARG ASN GLY TYR THR ILE LYS ASP ASP ILE
ILE ALA VAL LYS TRP GLU ASP SER PHE ALA GLU GLN GLY TYR LYS LYS TYR VAL SER ASN TYR LEU GLU SER ASN THR ALA GLY ASP PRO TYR THR ASP TYR GLU LYS ALA SER GLY SER PHE ASP LYS ALA ILE LYS THR GLU ALA ALA ARG ASP PRO LEU VAL ALA ALA TYR PRO ILE VAL
GLY VAL GLY MET GLU LYS LEU ILE ILE SER THR ASN GLU HIS ALA SER THR 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 ILE ASN LYS GLY GLU SER ALA TYR ILE ASN ALA ASN VAL THR GLY GLN ASN GLY PHE THR THR ASN LEU VAL LEU ASP GLY ASP THR LEU SER
THR ILE LYS ALA GLN GLN ILE GLY ASN GLN ASP THR TYR PRO LYS LYS GLY LEU SER PRO LEU ALA LEU ASN THR MET ASP GLN PHE SER SER ARG LEU ILE PRO ILE ASN TYR ASP GLN LEU LYS LYS LEU ASP ALA GLY LYS GLN ILE LYS LEU GLU THR
THR GLN VAL SER GLY ASN PHE GLY THR LYS ASN SER SER GLY GLN ILE VAL THR GLU GLY ASN SER TRP SER ASP TYR ILE SER GLN ILE ASP SER ILE SER ALA ILE ALA GLU ASP THR GLU ASN GLU SER TYR GLU ARG ARG VAL THR ALA LYS ASN LEU GLN ILE ASP PRO
LYS THR PRO GLU LEU THR ILE GLY GLU ALA ILE GLU LYS ALA PHE GLY ALA THR LYS LYS ASP GLY LEU LEU TYR PHE ASN ASP ILE PRO ILE ASP GLU SER CYS VAL GLU LEU ILE PHE ASP ASP ASN THR ALA ASN LYS ILE LYS ASP SER LEU LYS THR LEU SER ASP LYS LYS ILE
TYR ASN VAL LYS LEU GLU ARG GLY MET ASN ILE LEU ILE LYS THR 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 SER ILE THR GLU LEU ASN SER THR PRO GLU ILE LEU ASP GLU PRO GLU VAL LYS ILE PRO THR D758 I761 M762 H765 D771 S777 T781 P790 T791 N794 D799 Y800 I801 Q802 K803
V806 E807 Q811 Y812 S813 D817 E824 D831 K836 T837 R843 E850 T854 K857 L858 R859 I860 E869 L870 L871 V872 LEU SER VAL ASP

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	13266	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 ($e^-/Å^2$)	110.1	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	0.100	Depositor
Minimum map value	-0.041	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.035	Depositor
Map size (Å)	362.5, 362.5, 362.5	wwPDB
Map dimensions	250, 250, 250	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.45, 1.45, 1.45	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	0.44	0/967	0.62	0/1307
1	B	0.44	0/967	0.62	0/1307
1	C	0.44	0/967	0.62	0/1307
1	D	0.44	0/967	0.62	0/1307
1	E	0.44	0/967	0.62	0/1307
1	F	0.44	0/967	0.62	0/1307
1	G	0.44	0/967	0.62	0/1307
1	H	0.44	0/967	0.62	0/1307
1	I	0.44	0/967	0.62	0/1307
1	J	0.44	0/967	0.62	0/1307
1	K	0.44	0/967	0.62	0/1307
1	L	0.44	0/967	0.62	0/1307
1	M	0.44	0/967	0.62	0/1307
1	N	0.44	0/967	0.62	0/1307
All	All	0.44	0/13538	0.62	0/18298

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

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	946	0	914	11	0
1	B	946	0	914	11	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	946	0	914	11	0
1	D	946	0	914	9	0
1	E	946	0	914	9	0
1	F	946	0	914	9	0
1	G	946	0	914	8	0
1	H	946	0	914	10	0
1	I	946	0	914	9	0
1	J	946	0	914	11	0
1	K	946	0	914	11	0
1	L	946	0	914	11	0
1	M	946	0	914	12	0
1	N	946	0	914	8	0
All	All	13244	0	12796	139	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 139 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:F:858:LEU:HG	1:F:860:ILE:HD11	1.78	0.66
1:I:858:LEU:HG	1:I:860:ILE:HD11	1.78	0.66
1:J:858:LEU:HG	1:J:860:ILE:HD11	1.78	0.66
1:M:858:LEU:HG	1:M:860:ILE:HD11	1.78	0.66
1:E:858:LEU:HG	1:E:860:ILE:HD11	1.78	0.65

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	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	B	113/876 (13%)	93 (82%)	20 (18%)	0	100	100
1	C	113/876 (13%)	93 (82%)	20 (18%)	0	100	100
1	D	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	E	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	F	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	G	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	H	113/876 (13%)	93 (82%)	20 (18%)	0	100	100
1	I	113/876 (13%)	93 (82%)	20 (18%)	0	100	100
1	J	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	K	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	L	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	M	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
1	N	113/876 (13%)	92 (81%)	21 (19%)	0	100	100
All	All	1582/12264 (13%)	1292 (82%)	290 (18%)	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	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	B	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	C	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	D	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	E	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	F	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	G	102/790 (13%)	93 (91%)	9 (9%)	10	34

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	H	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	I	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	J	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	K	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	L	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	M	102/790 (13%)	93 (91%)	9 (9%)	10	34
1	N	102/790 (13%)	93 (91%)	9 (9%)	10	34
All	All	1428/11060 (13%)	1302 (91%)	126 (9%)	13	34

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	781	THR
1	M	777	SER
1	H	872	VAL
1	M	761	ILE
1	N	761	ILE

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.

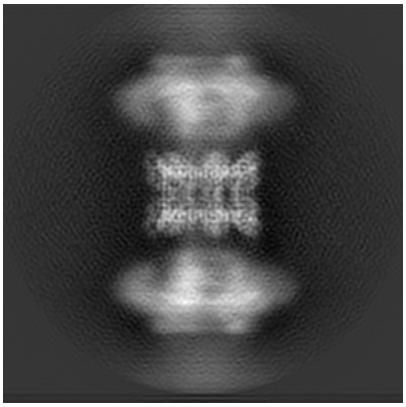
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-20103. These allow visual inspection of the internal detail of the map and identification of artifacts.

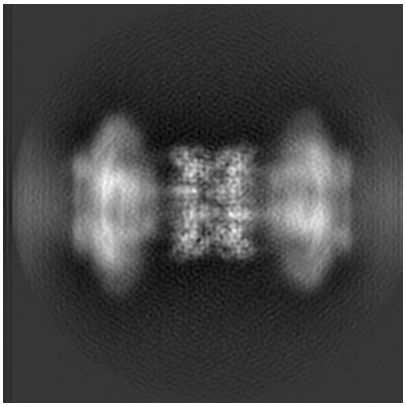
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

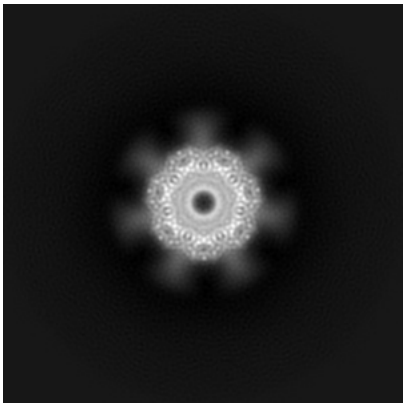
6.1.1 Primary map



X



Y

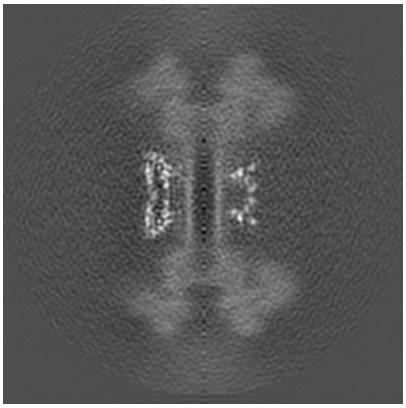


Z

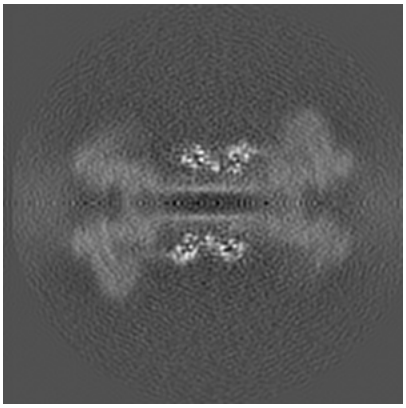
The images above show the map projected in three orthogonal directions.

6.2 Central slices

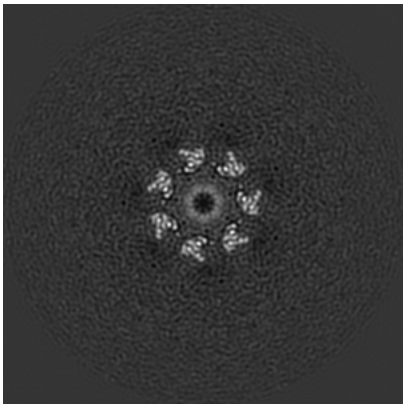
6.2.1 Primary map



X Index: 125



Y Index: 125

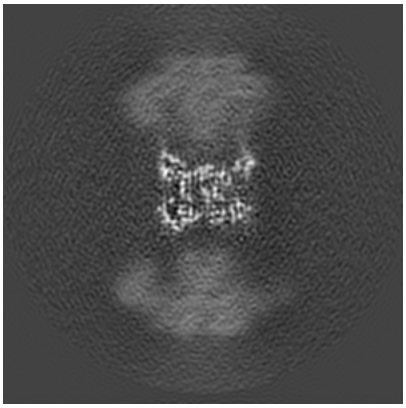


Z Index: 125

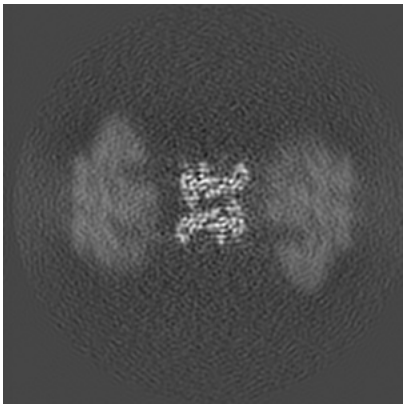
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

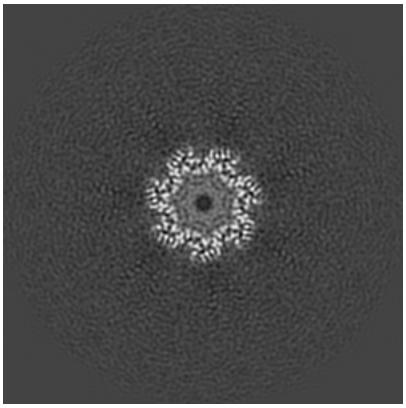
6.3.1 Primary map



X Index: 146



Y Index: 146

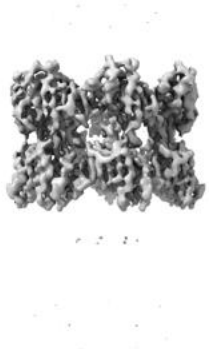


Z Index: 144

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

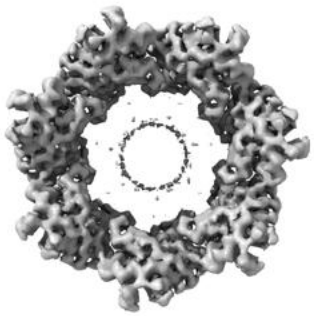
6.4 Orthogonal surface views (i)

6.4.1 Primary map



X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.035. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

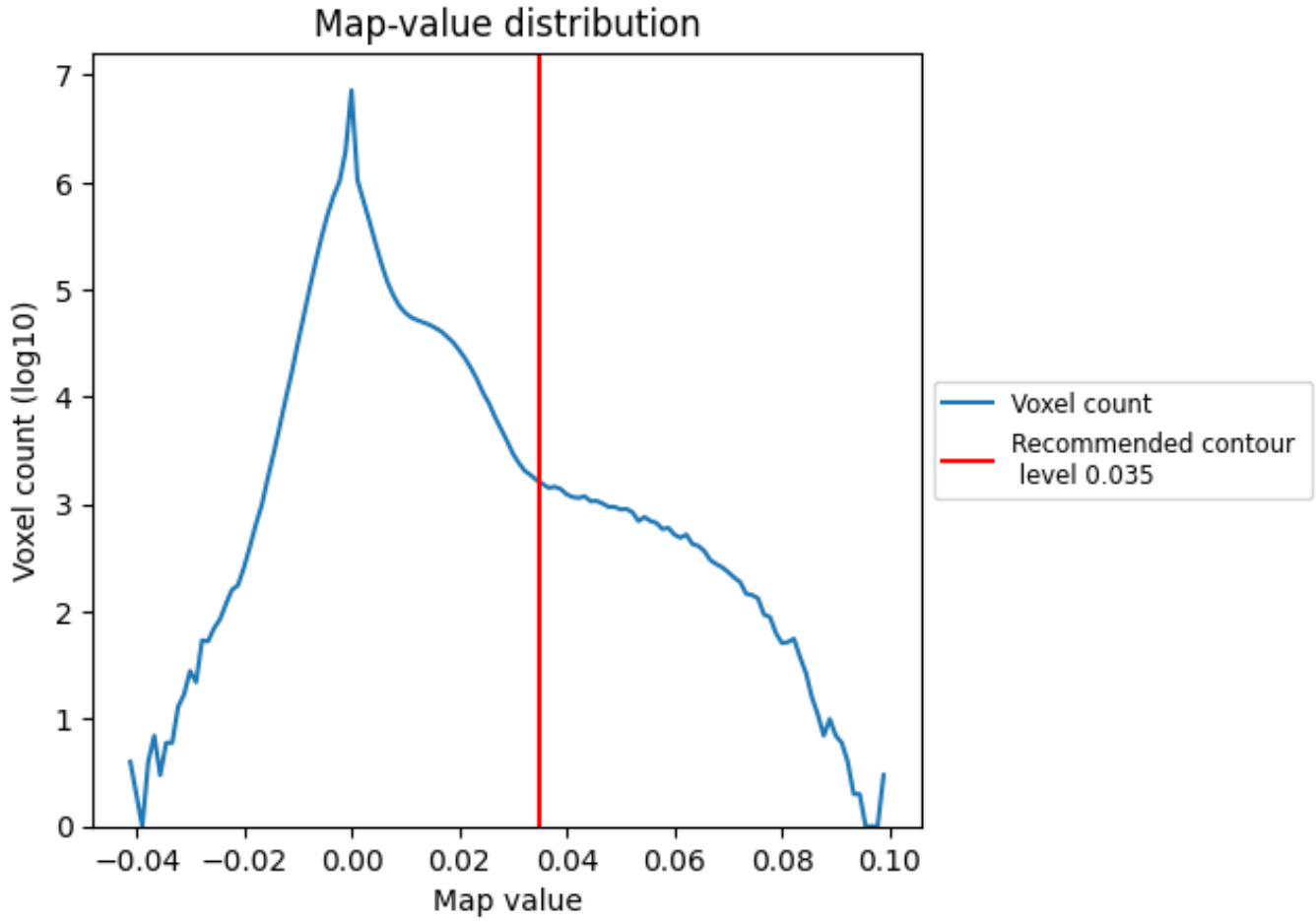
6.5 Mask visualisation (i)

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

7 Map analysis

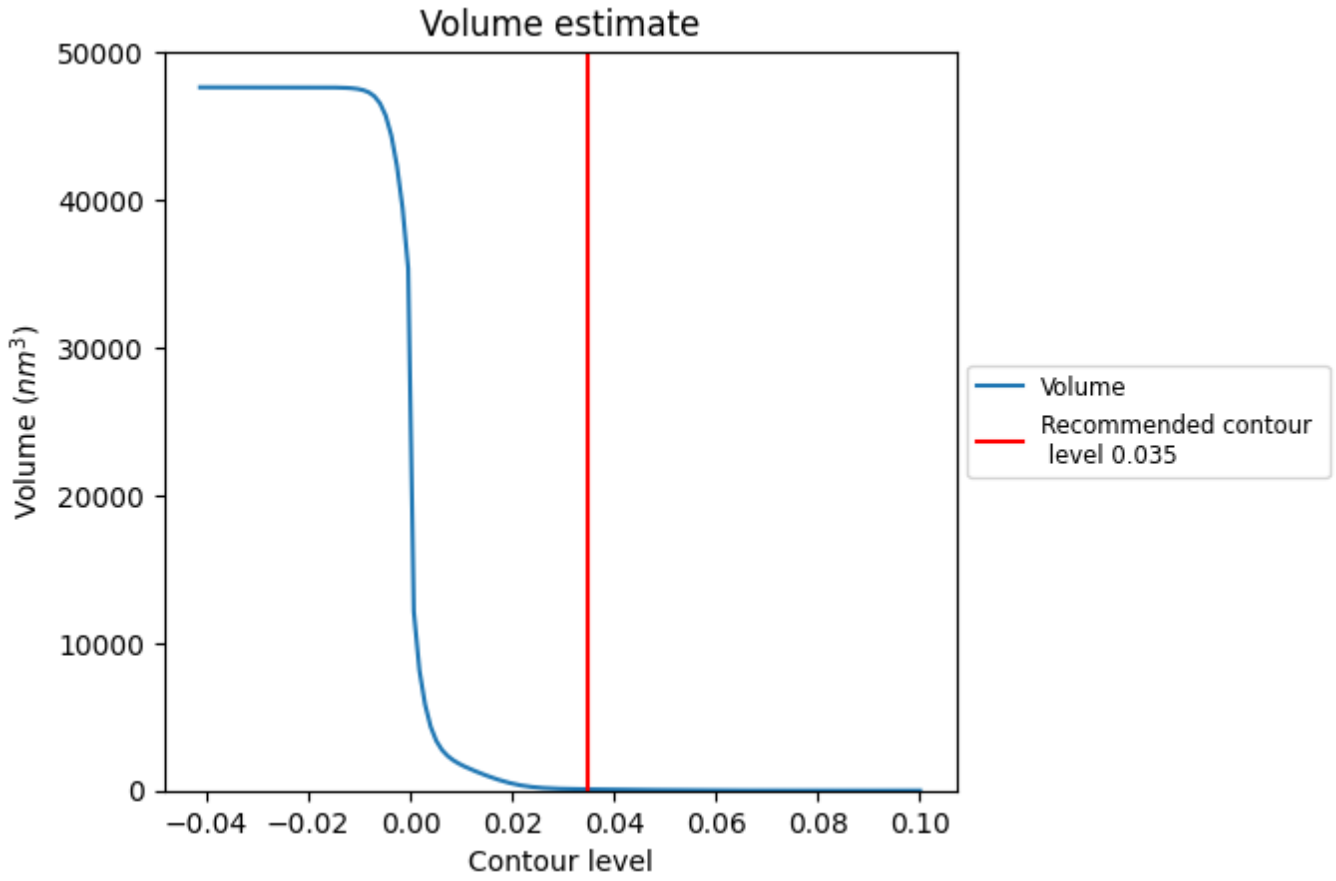
This section contains the results of statistical analysis of the map.

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.

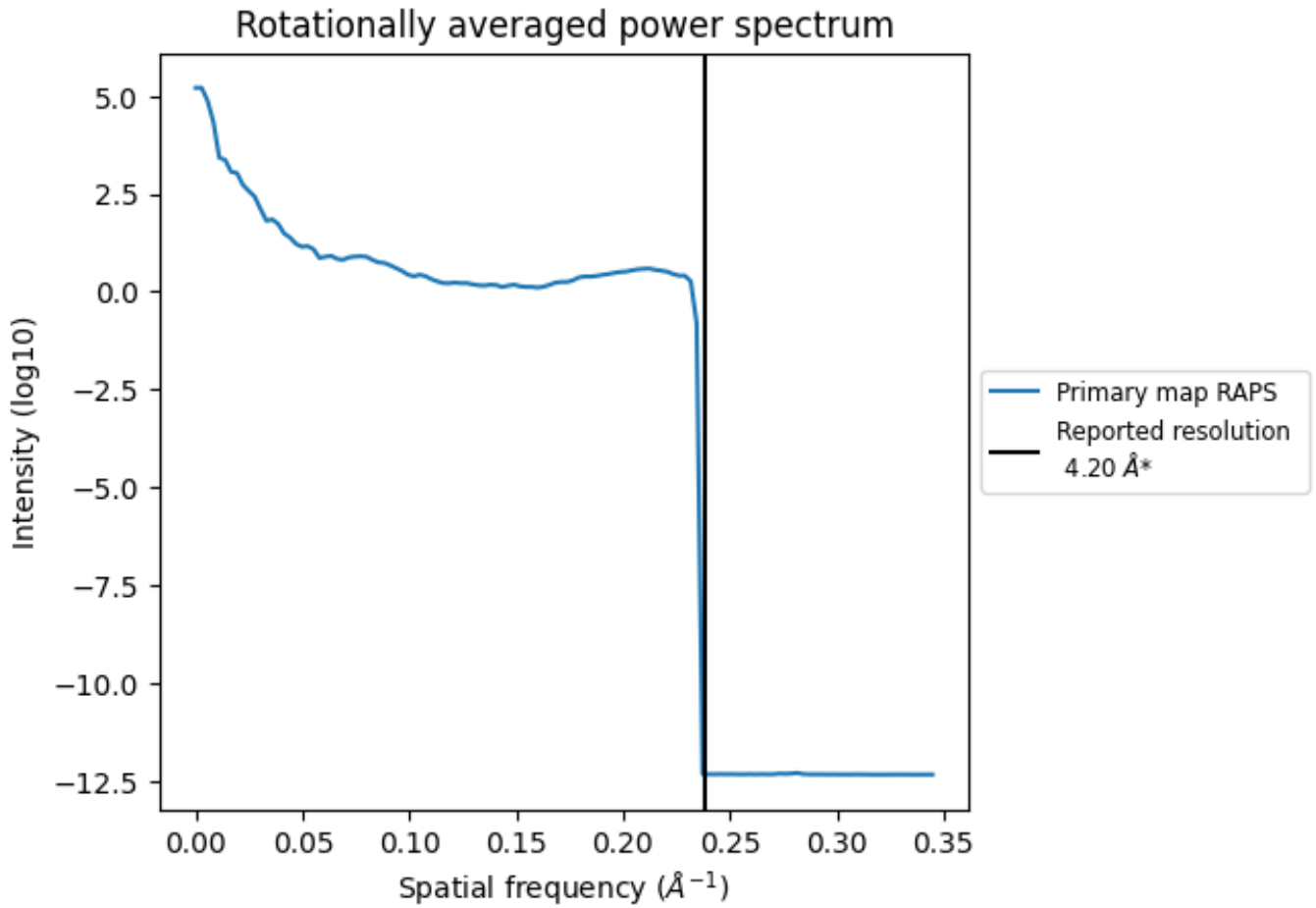
7.2 Volume estimate (i)



The volume at the recommended contour level is 86 nm^3; this corresponds to an approximate mass of 78 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)

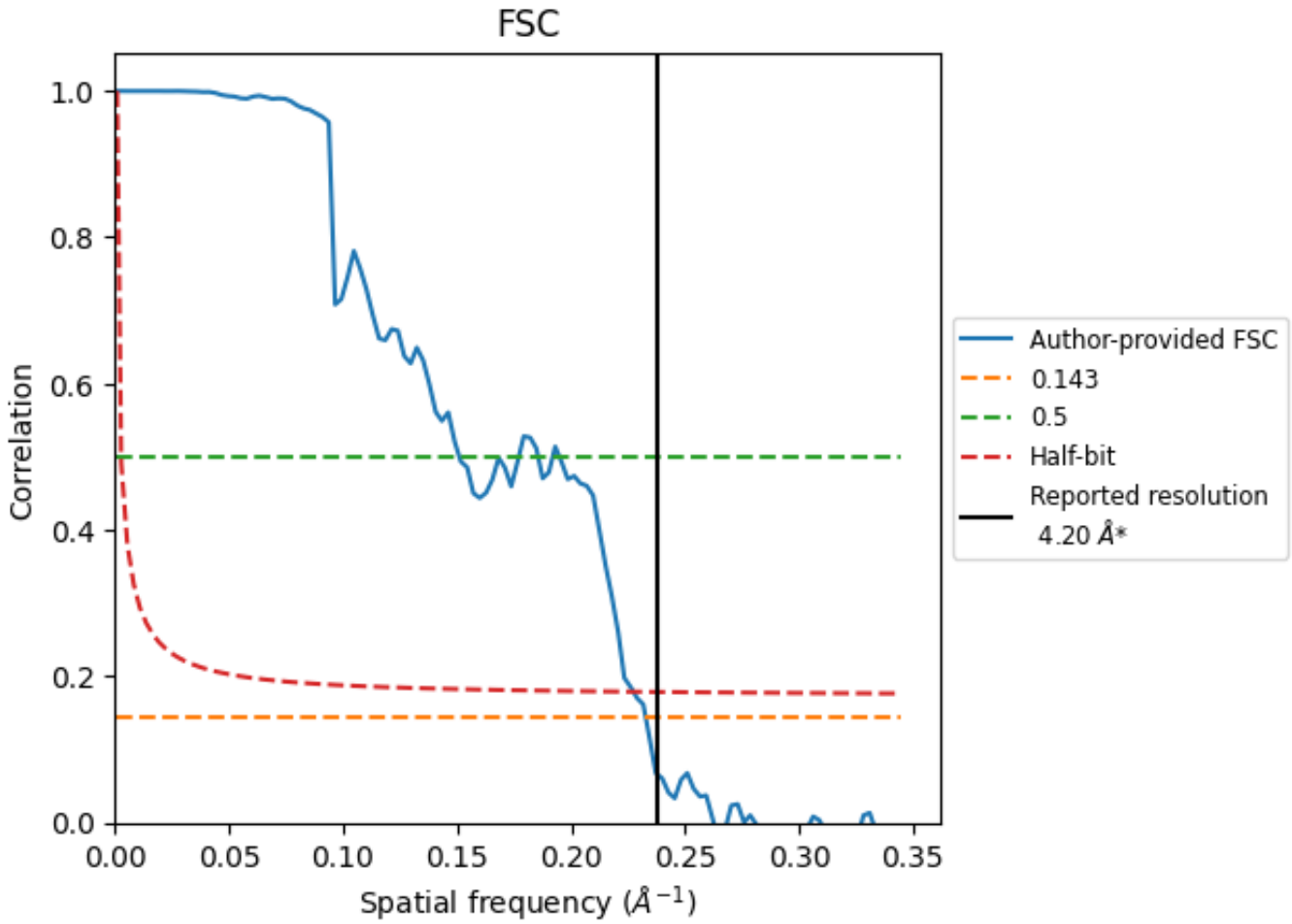


*Reported resolution corresponds to spatial frequency of 0.238 $Å^{-1}$

8 Fourier-Shell correlation

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 0.238 Å^{-1}

8.2 Resolution estimates (i)

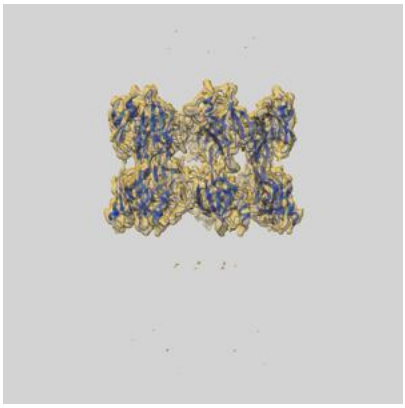
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.20	-	-
Author-provided FSC curve	4.30	6.62	4.40
Unmasked-calculated*	-	-	-

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

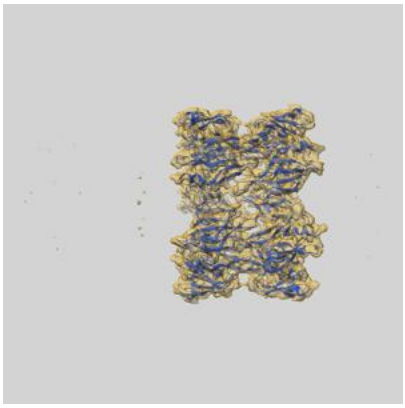
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-20103 and PDB model 6OKS. Per-residue inclusion information can be found in section 3 on page 5.

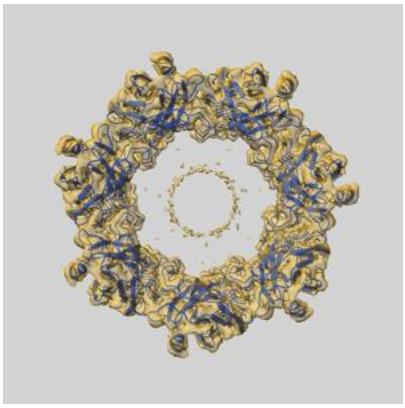
9.1 Map-model overlay



X



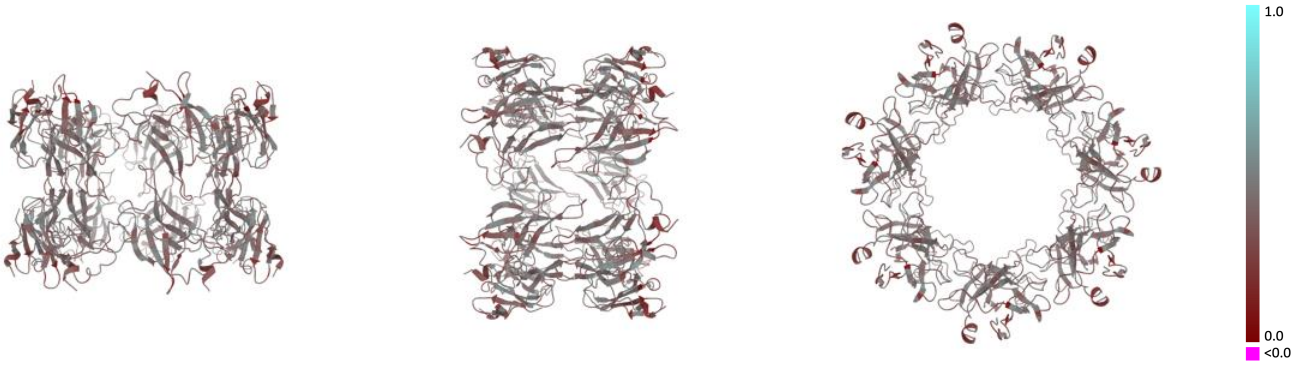
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.035 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.

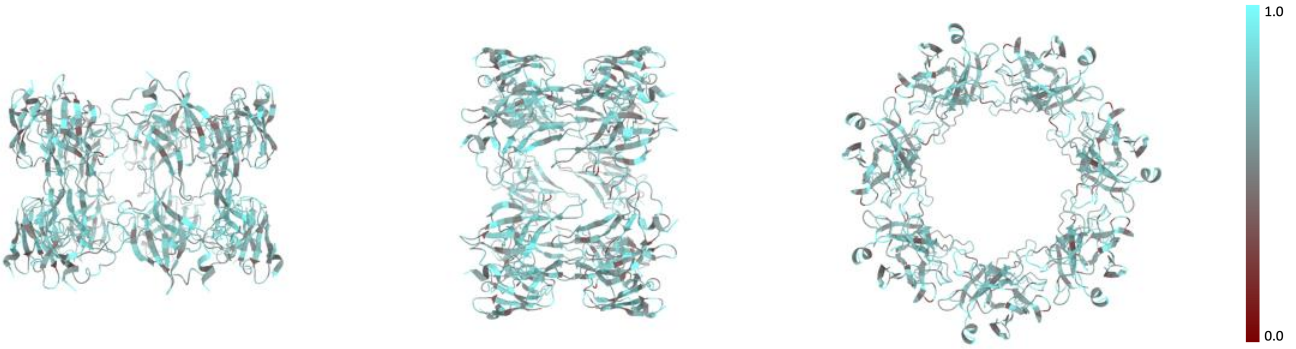
9.2 Q-score mapped to coordinate model (i)



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.

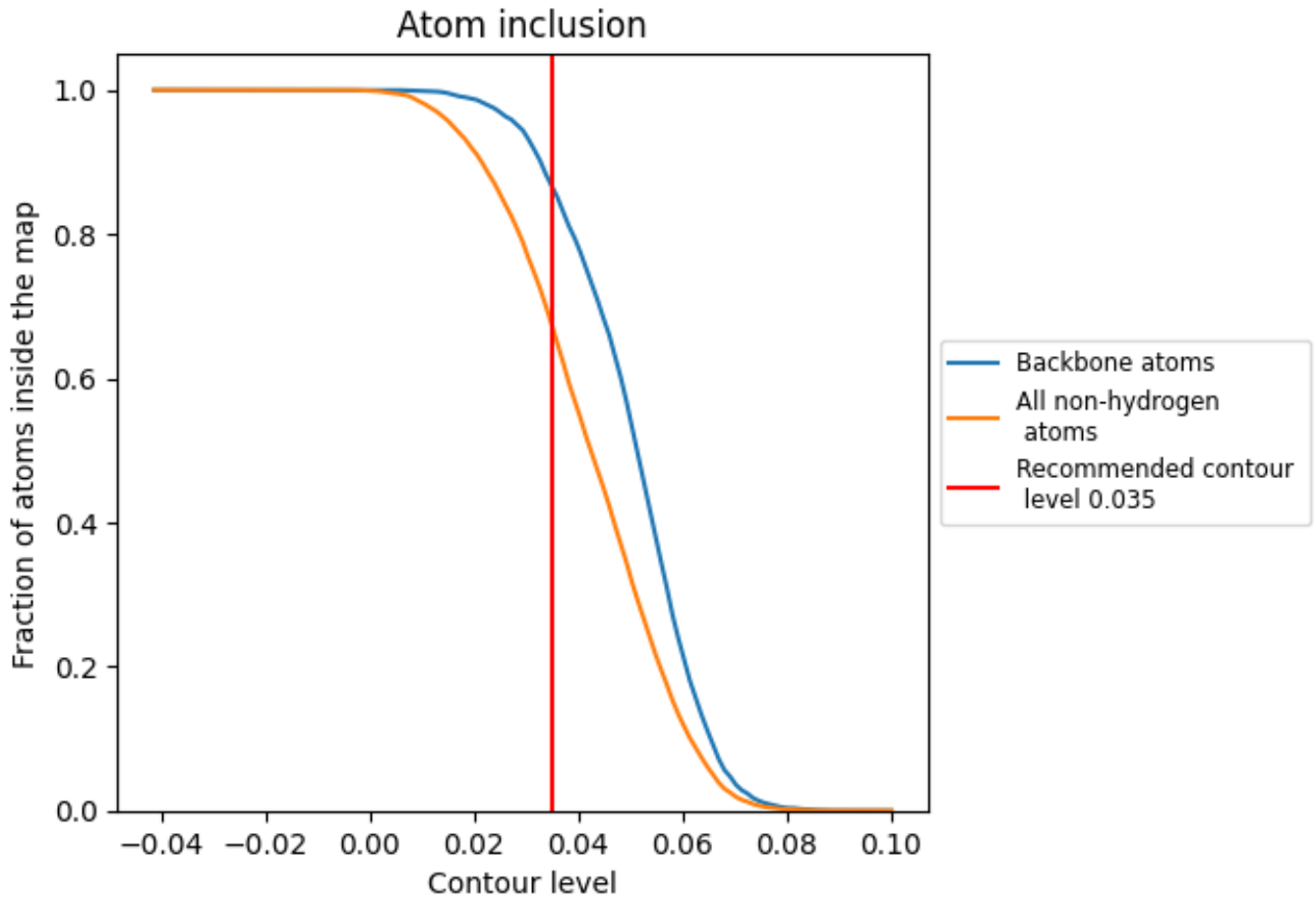
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.035).

9.4 Atom inclusion (i)



At the recommended contour level, 87% of all backbone atoms, 67% 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 (0.035) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6727	0.3890
A	0.6941	0.3960
B	0.6897	0.4000
C	0.6941	0.4020
D	0.6941	0.4010
E	0.6951	0.4010
F	0.6865	0.3970
G	0.6995	0.3980
H	0.6530	0.3780
I	0.6454	0.3800
J	0.6519	0.3800
K	0.6486	0.3810
L	0.6551	0.3810
M	0.6562	0.3780
N	0.6541	0.3800