wwPDB EM Validation Summary Report (i)

Mar 3, 2024 – 02:06 PM EST

PDB ID	:	5WLN
EMDB ID	:	EMD-8860
Title	:	Cryo-EM structure of the T2SS secretin XcpQ from Pseudomonas aeruginosa
Authors	:	Hay, I.D.; Belousoff, M.J.; Lithgow, T.J.
Deposited on	:	2017-07-27
Resolution	:	3.57 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:

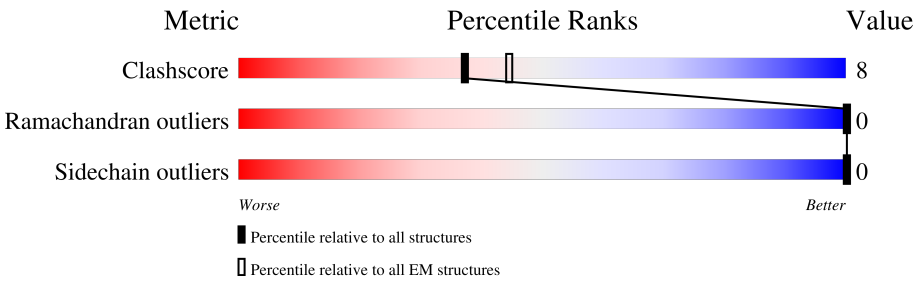
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.57 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	624	5% 52% 12% 37%
1	B	624	5% 51% 13% 37%
1	C	624	5% 49% 14% 37%
1	D	624	5% 49% 15% 37%
1	E	624	5% 51% 13% 37%
1	F	624	• 51% 12% 37%
1	G	624	• 51% 12% 37%
1	H	624	5% 52% 12% 37%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	624	5% 52% 11% 37%
1	J	624	5% 52% 11% 37%
1	K	624	5% 52% 11% 37%
1	L	624	5% 53% 11% 37%
1	M	624	5% 53% 11% 37%
1	N	624	5% 52% 12% 37%
1	O	624	6% 52% 12% 37%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 44415 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 Type II secretion system protein D.

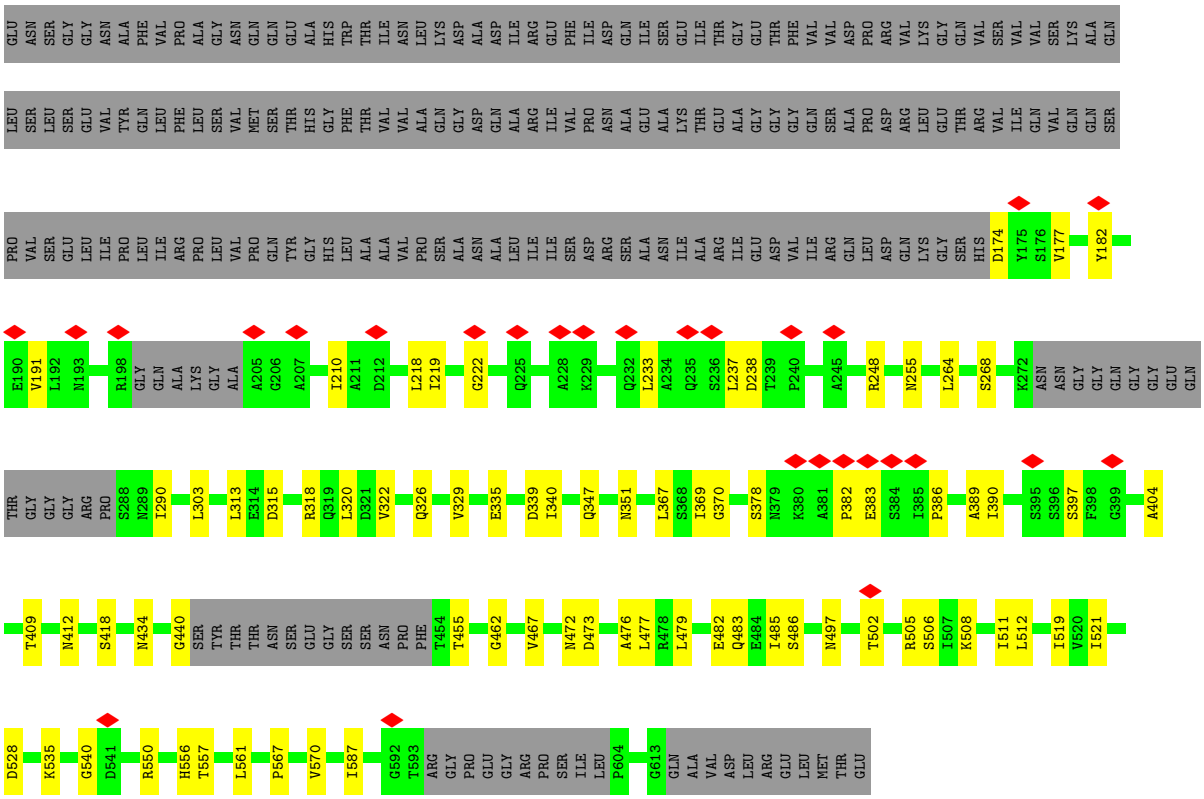
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	G	396	2961	1839	532	585	5	0	0
1	O	396	2961	1839	532	585	5	0	0
1	A	396	2961	1839	532	585	5	0	0
1	B	396	2961	1839	532	585	5	0	0
1	C	396	2961	1839	532	585	5	0	0
1	D	396	2961	1839	532	585	5	0	0
1	E	396	2961	1839	532	585	5	0	0
1	F	396	2961	1839	532	585	5	0	0
1	H	396	2961	1839	532	585	5	0	0
1	I	396	2961	1839	532	585	5	0	0
1	J	396	2961	1839	532	585	5	0	0
1	K	396	2961	1839	532	585	5	0	0
1	L	396	2961	1839	532	585	5	0	0
1	M	396	2961	1839	532	585	5	0	0
1	N	396	2961	1839	532	585	5	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

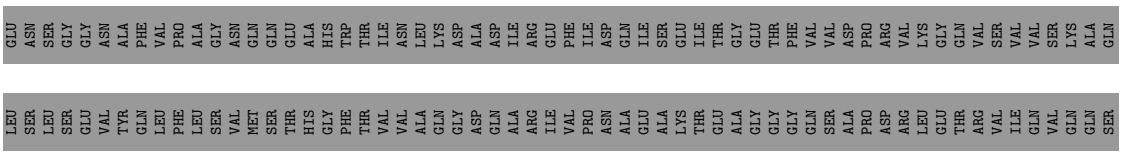
- Molecule 1: Type II secretion system protein D

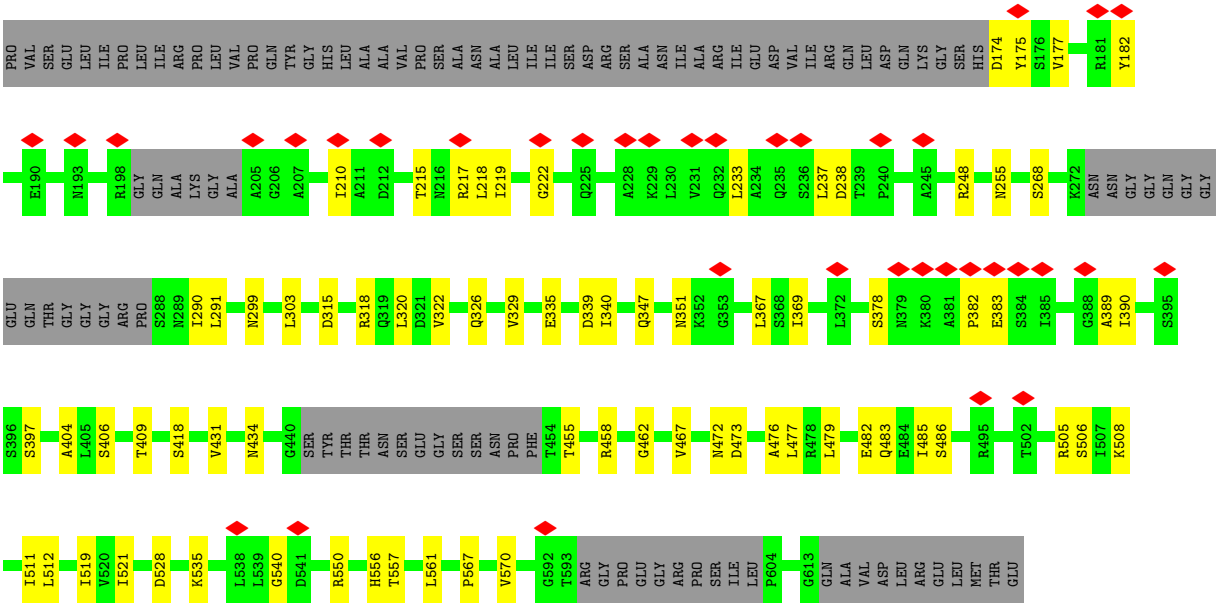
Chain G:



- Molecule 1: Type II secretion system protein D

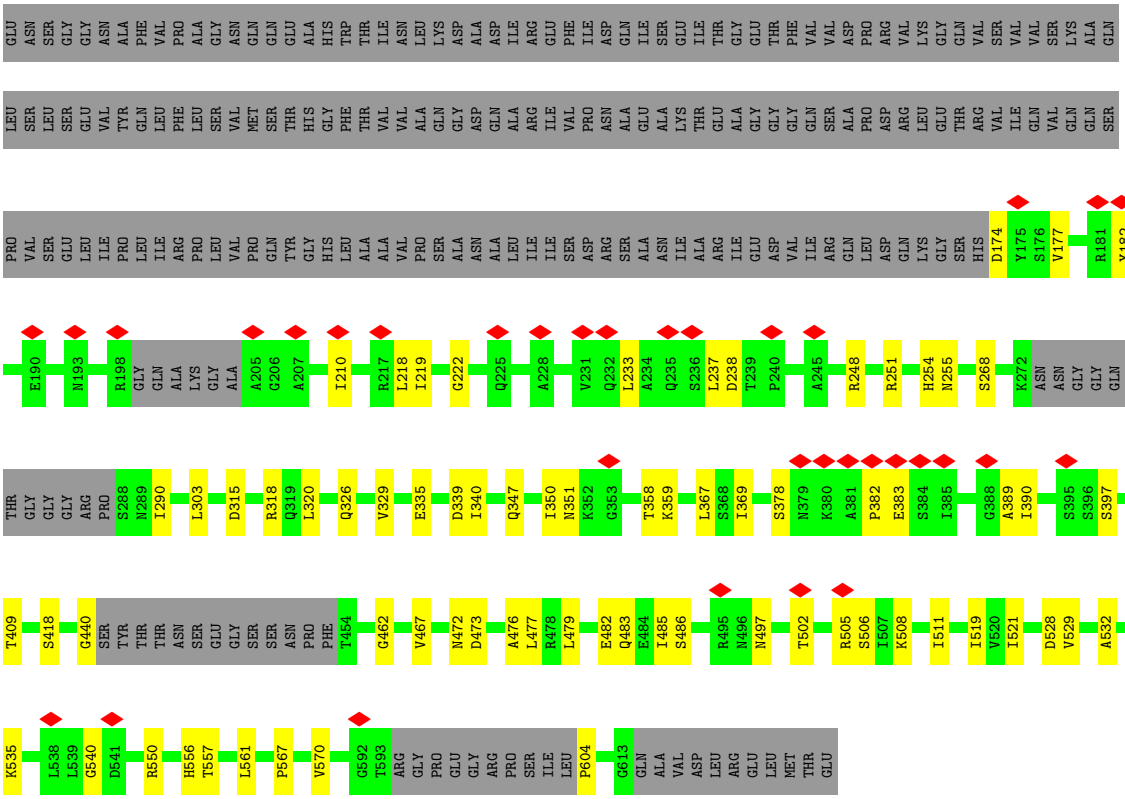
Chain O:





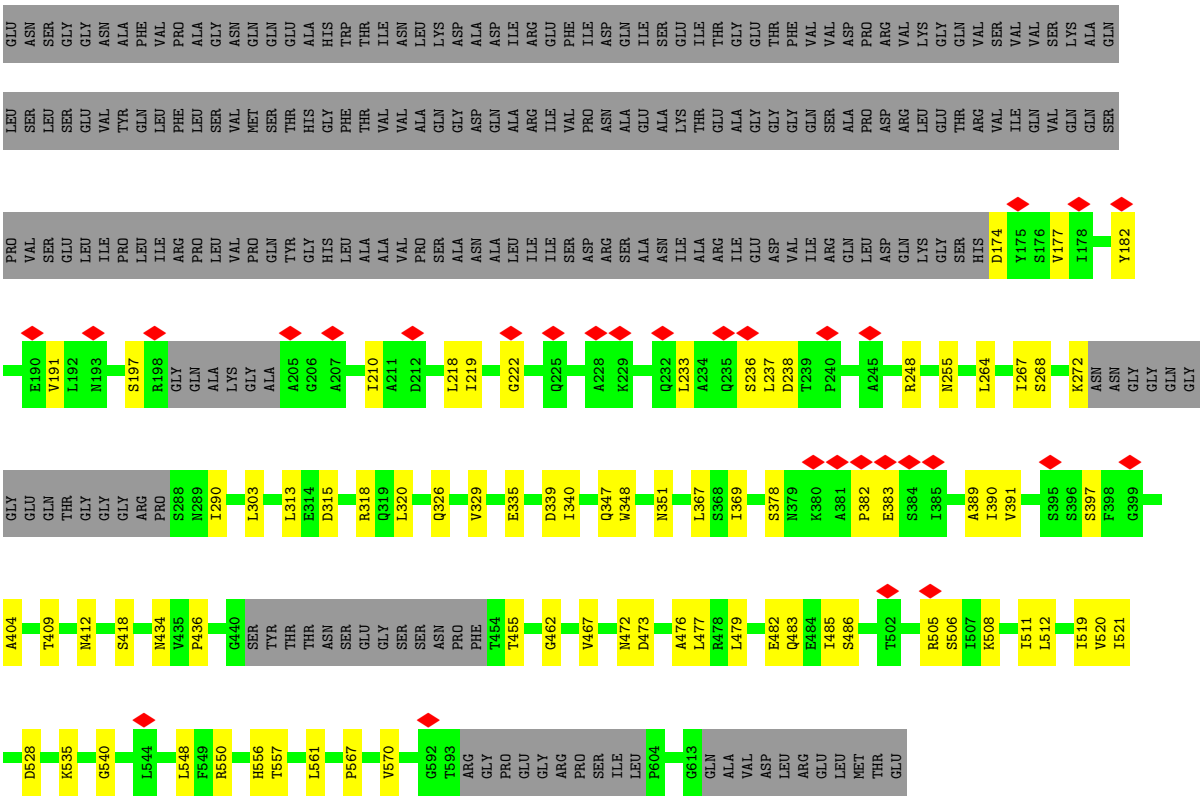
- Molecule 1: Type II secretion system protein D

Chain A:



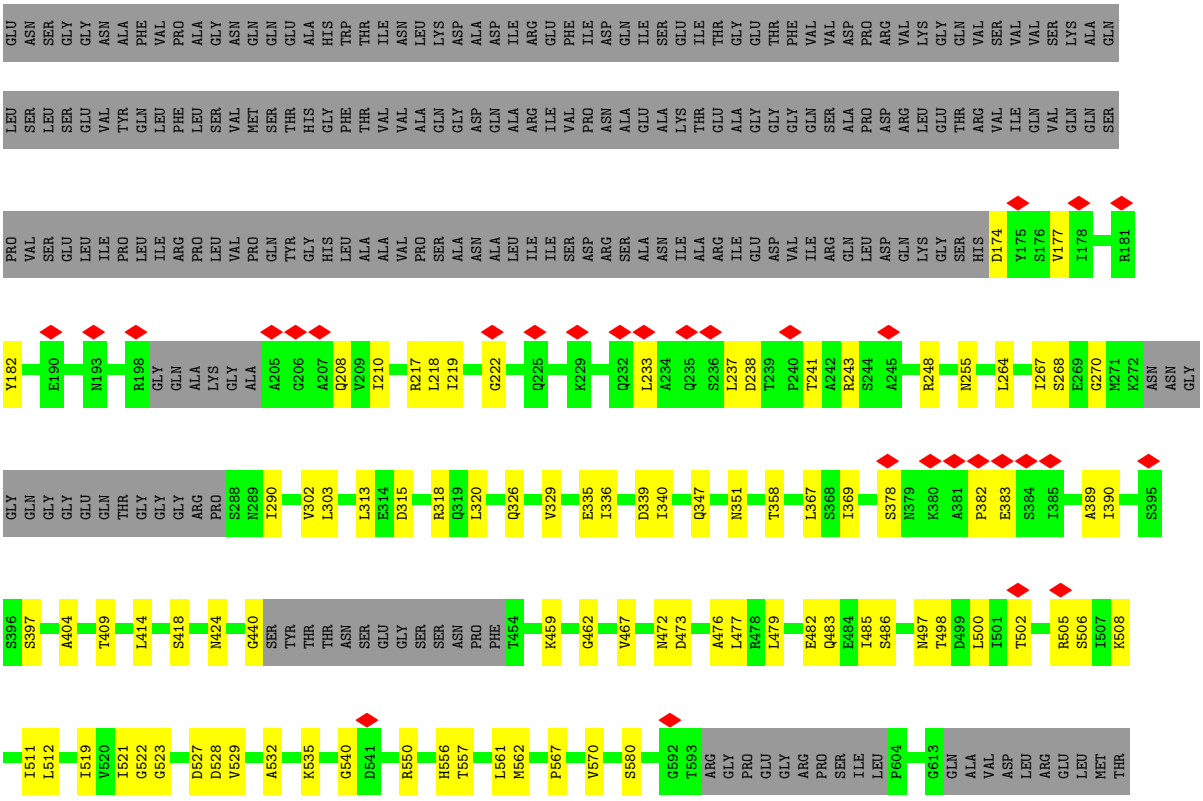
- Molecule 1: Type II secretion system protein D

Chain B:



- Molecule 1: Type II secretion system protein D

Chain C: 5% 49% 14% 37%



GLU

• Molecule 1: Type II secretion system protein D

Chain D:

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

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

PRO VAL SER GLU LEU ILE PRO LEU ILE ARG PRO LEU VAL PRO GLN TYR GLY HIS LEU ALA ALA VAL PRO SER ALA ASN ALA LEU ILE ILE SER ASP ARG SER ALA ASN ILE ALA ARG ILE GLU ASP VAL ILE ARG GLN LEU ASP GLN LYS GLY SER HIS D174 Y175 S176 V177 R181 Y182

E190 N193 R198 GLY GLN ALA LYS GLY ALA A205 G206 A207 I210 A211 D212 A213 R214 T215 L218 I219 G222 Q225 A228 V231 Q232 L233 A234 Q235 S236 L237 D238 T239 P240 R243 S244 A245 R248 H254 N255 L264 I267 S268 K272 ASN ASN

GLY GLY GLN GLY GLY GLU GLN THR GLY GLY GLY ARG PRO S288 N289 I290 L291 N299 L303 L313 E314 D315 R318 Q319 L320 Q326 V329 E335 D339 I340 Q347 N351 M355 N361 L367 S368 I369 L373 Q374 S378 N379 K380 A381 P382 E383

S384 I385 P386 A389 I390 S395 S396 S397 F398 G399 A404 L405 S406 T409 N412 L413 L414 S418 L419 L420 N434 V435 P436 G440 SER TYR THR THR ASN SER GLU GLY SER SER ASN PRO PHE T454 T455 R458 G462 V467 N472 D473 A476 L477 R478

L479 E482 Q483 E484 I485 S486 T502 R505 S506 I507 K508 I511 L512 I519 V520 I521 D528 K535 L538 L539 G540 D541 L544 R550 H556 T557 L561 F564 L565 R566 P567 V570 K583 I587 G592 T593 ARG GLY PRO GLU GLY ARG PRO

SER ILE LEU P604 F611 D612 G613 GLN ALA VAL ASP LEU ARG GLU LEU MET THR GLU

• Molecule 1: Type II secretion system protein D

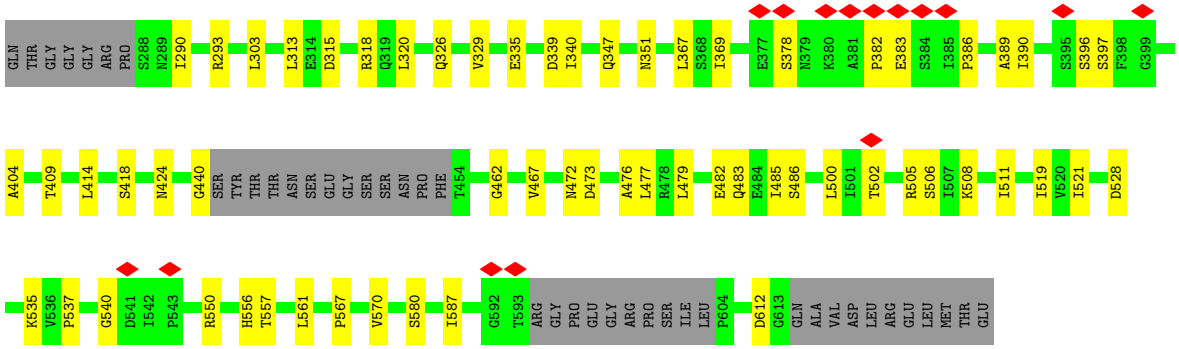
Chain E:

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

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

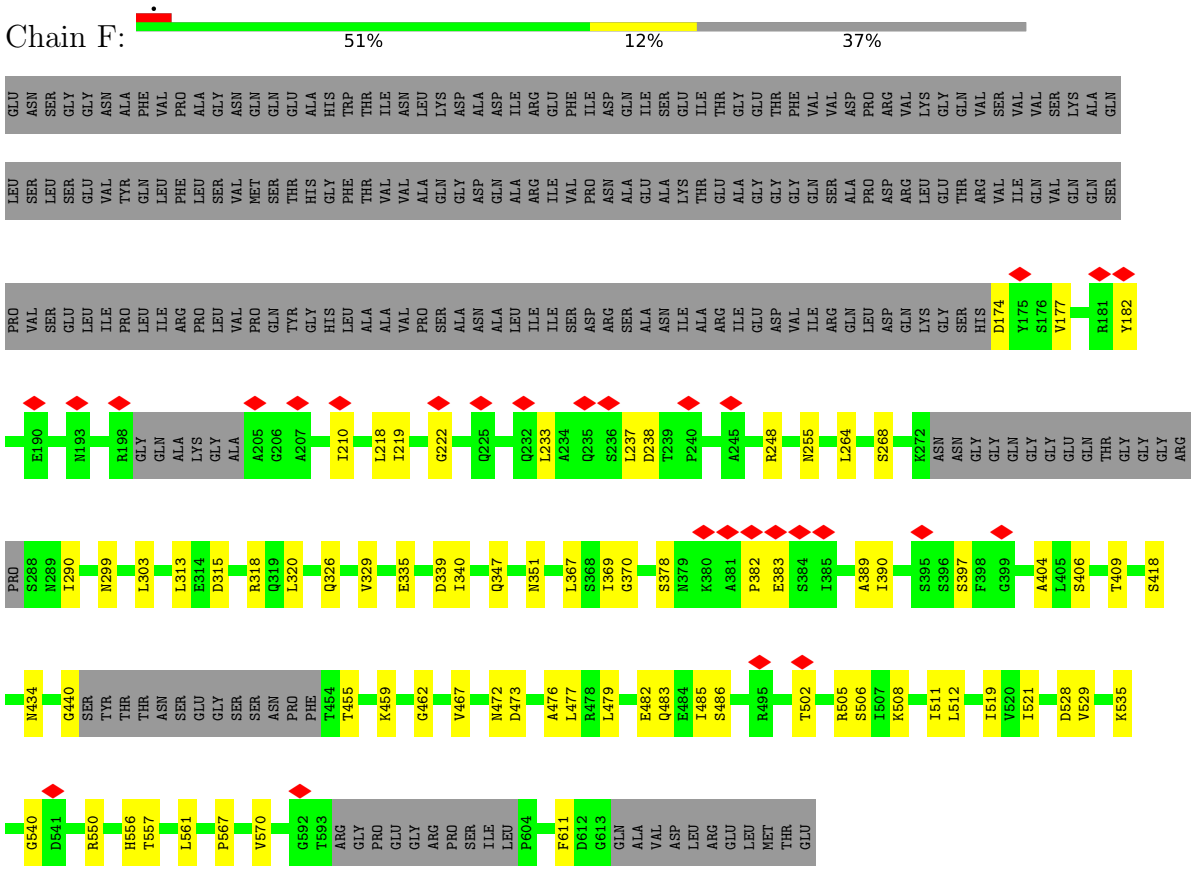
PRO VAL SER GLU LEU ILE PRO LEU ILE ARG PRO LEU VAL PRO GLN TYR GLY HIS LEU ALA ALA VAL PRO SER ALA ASN ALA LEU ILE ILE SER ASP ARG SER ALA ASN ILE ALA ARG ILE GLU ASP VAL ILE ARG GLN LEU ASP GLN LYS GLY SER HIS D174 Y175 S176 V177 I178 R181

Y182 E190 V191 L192 N193 R198 GLY GLN ALA LYS GLY ALA A205 G206 I210 R214 T215 N216 R217 L218 I219 G222 Q225 A226 K229 Q232 L233 A234 Q235 S236 L237 D238 A245 R248 N255 L264 S268 K272 ASN ASN GLY GLY GLN GLY GLY GLU



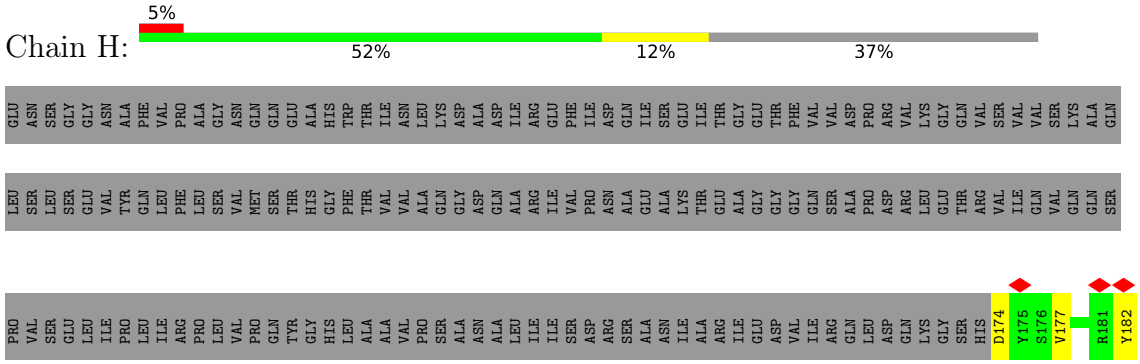
- Molecule 1: Type II secretion system protein D

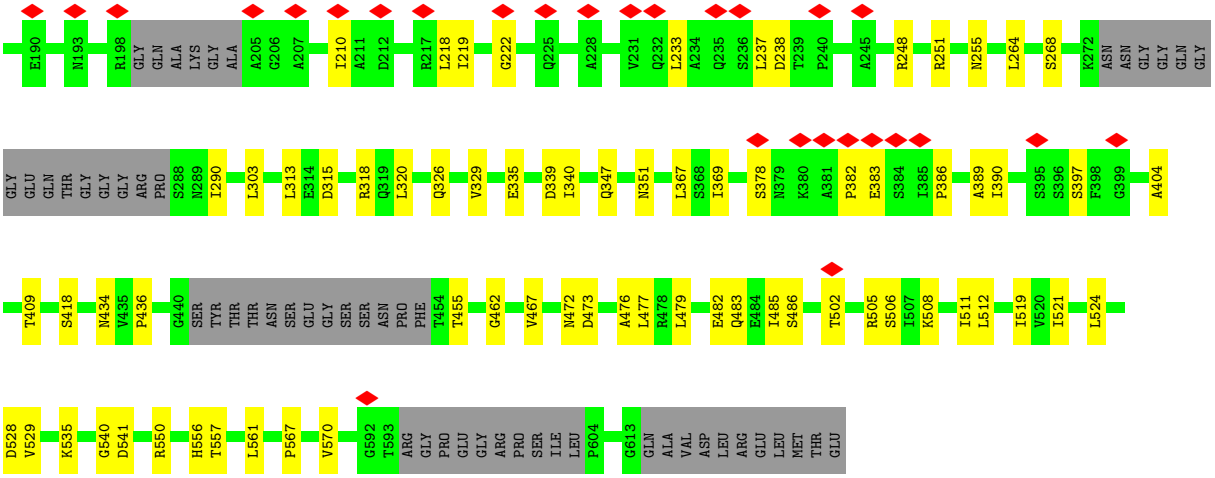
Chain F:



- Molecule 1: Type II secretion system protein D

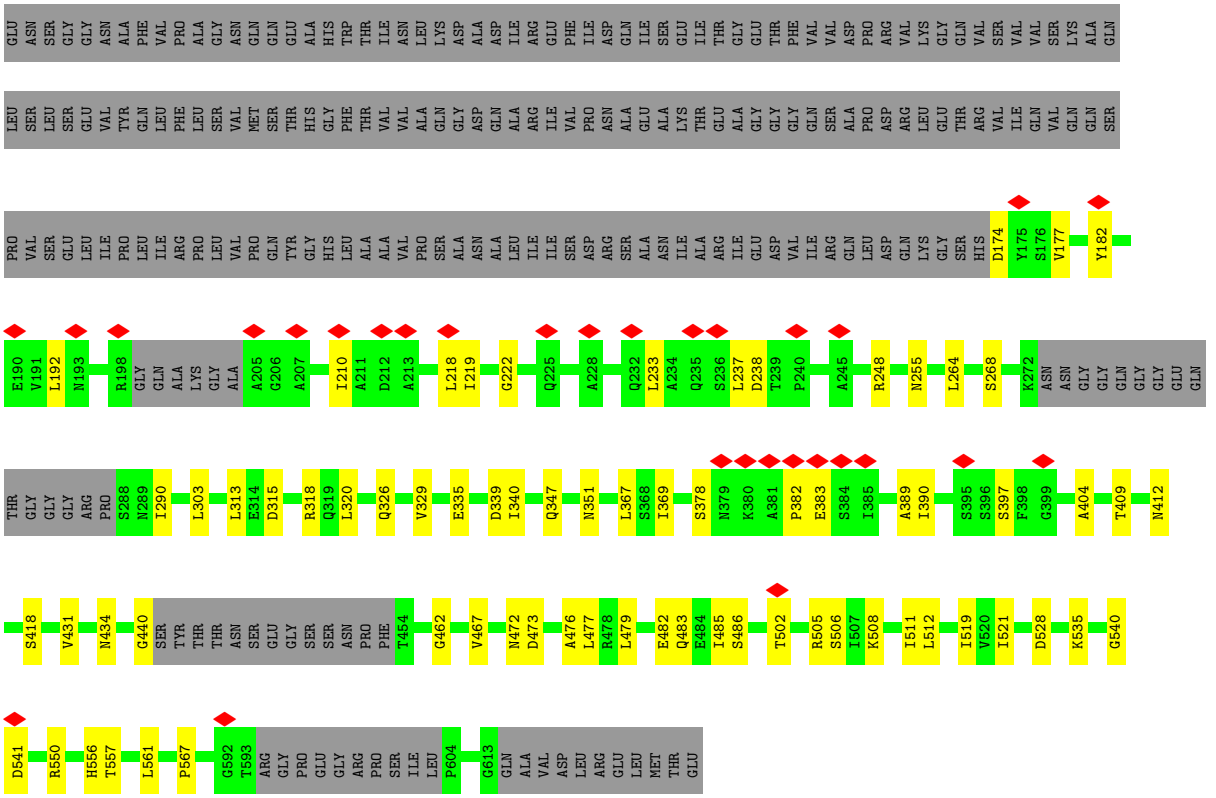
Chain H:





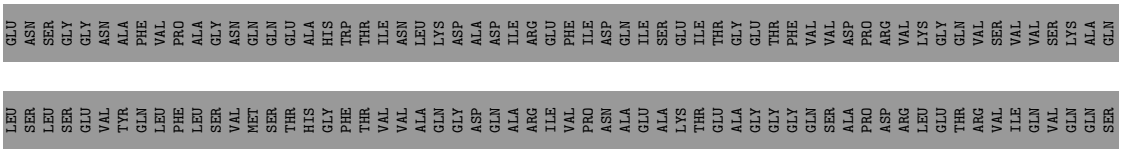
- Molecule 1: Type II secretion system protein D

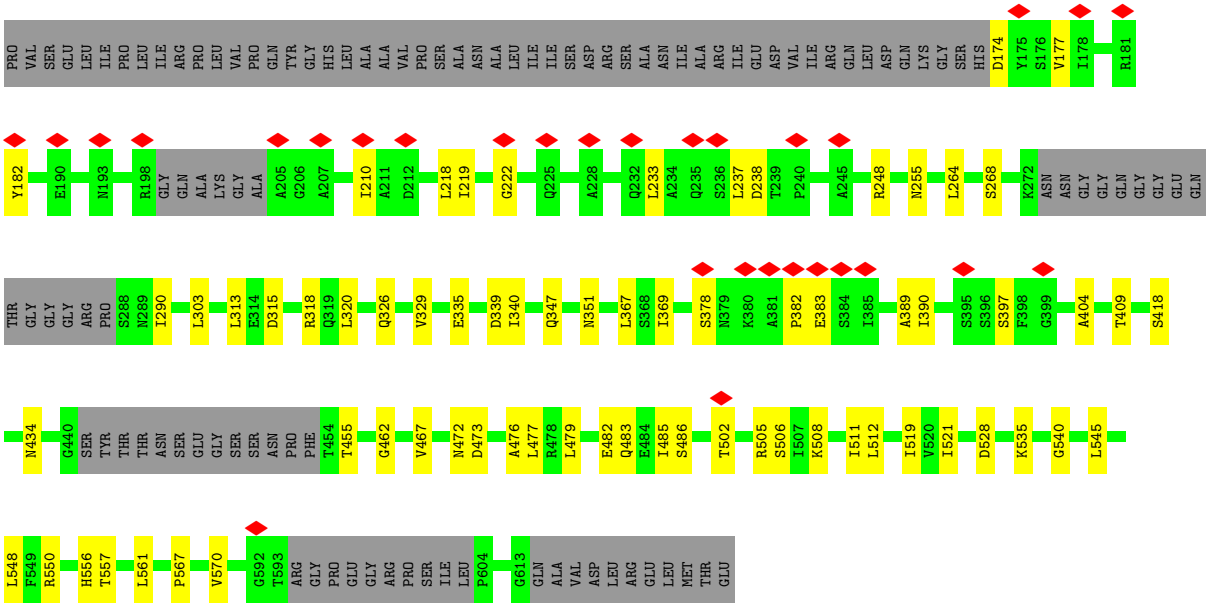
Chain I:



- Molecule 1: Type II secretion system protein D

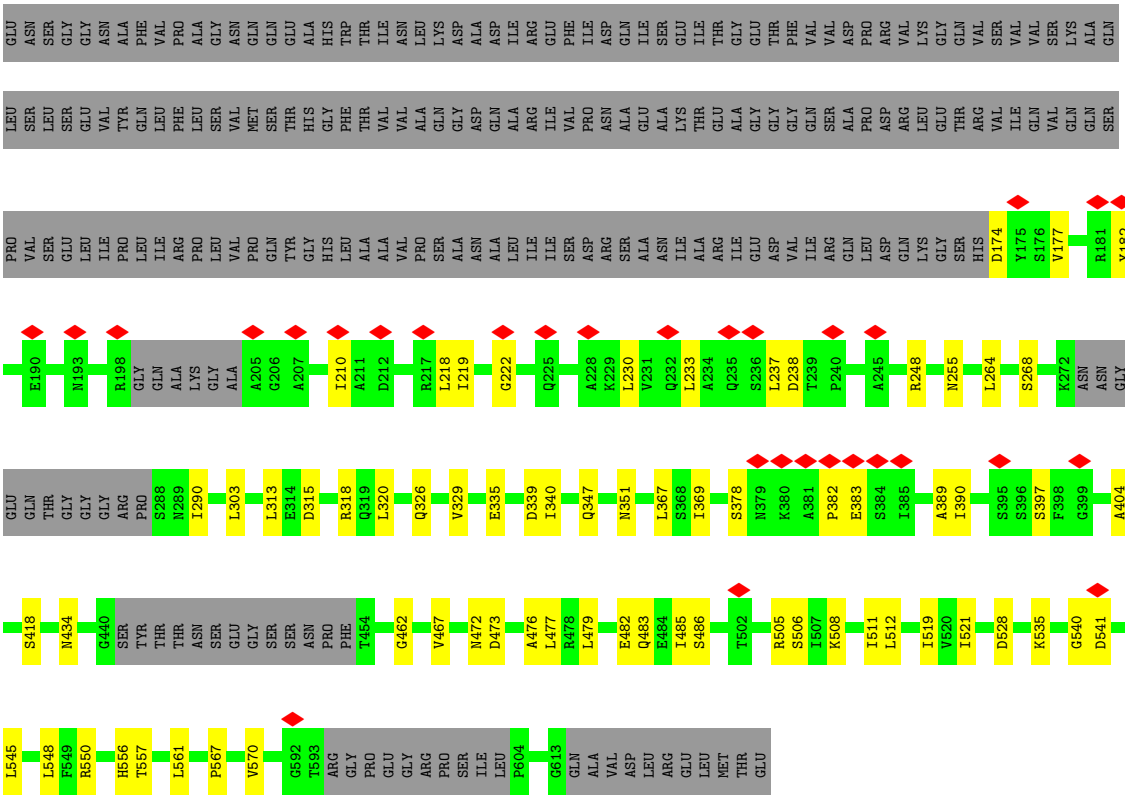
Chain J:





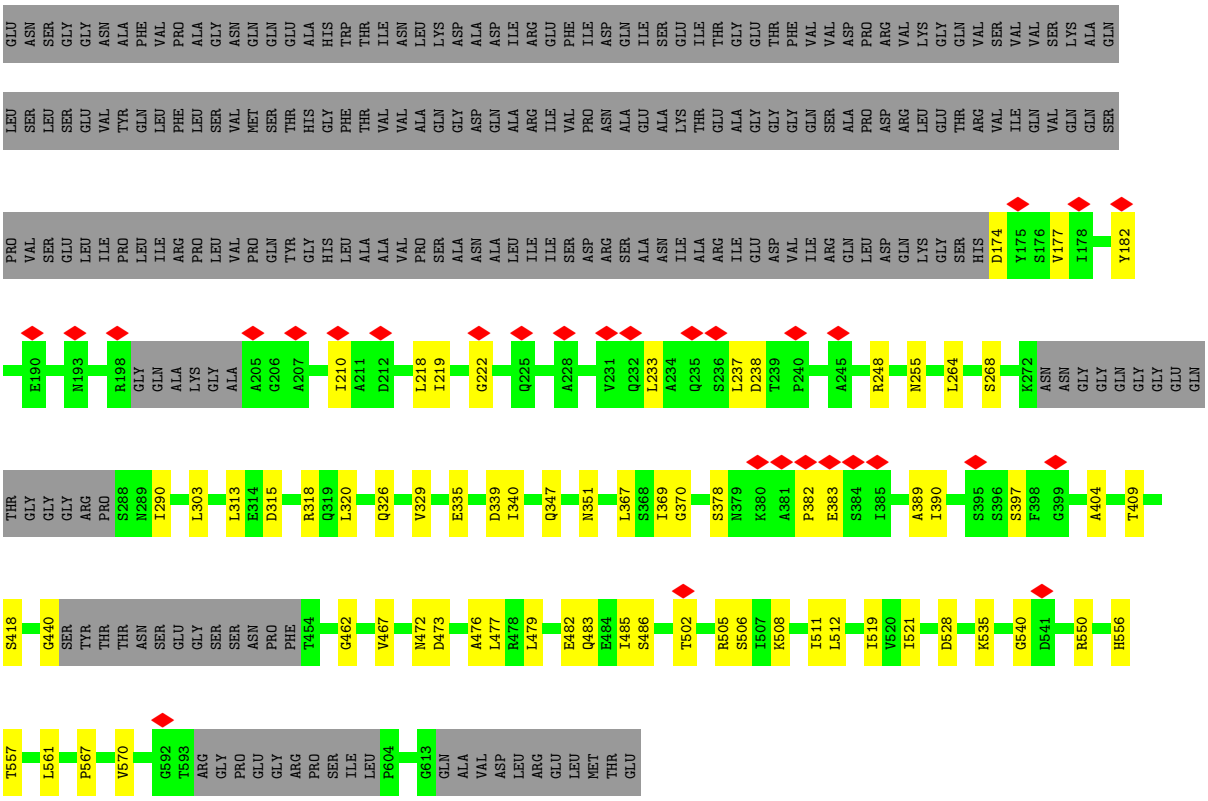
• Molecule 1: Type II secretion system protein D

Chain K:



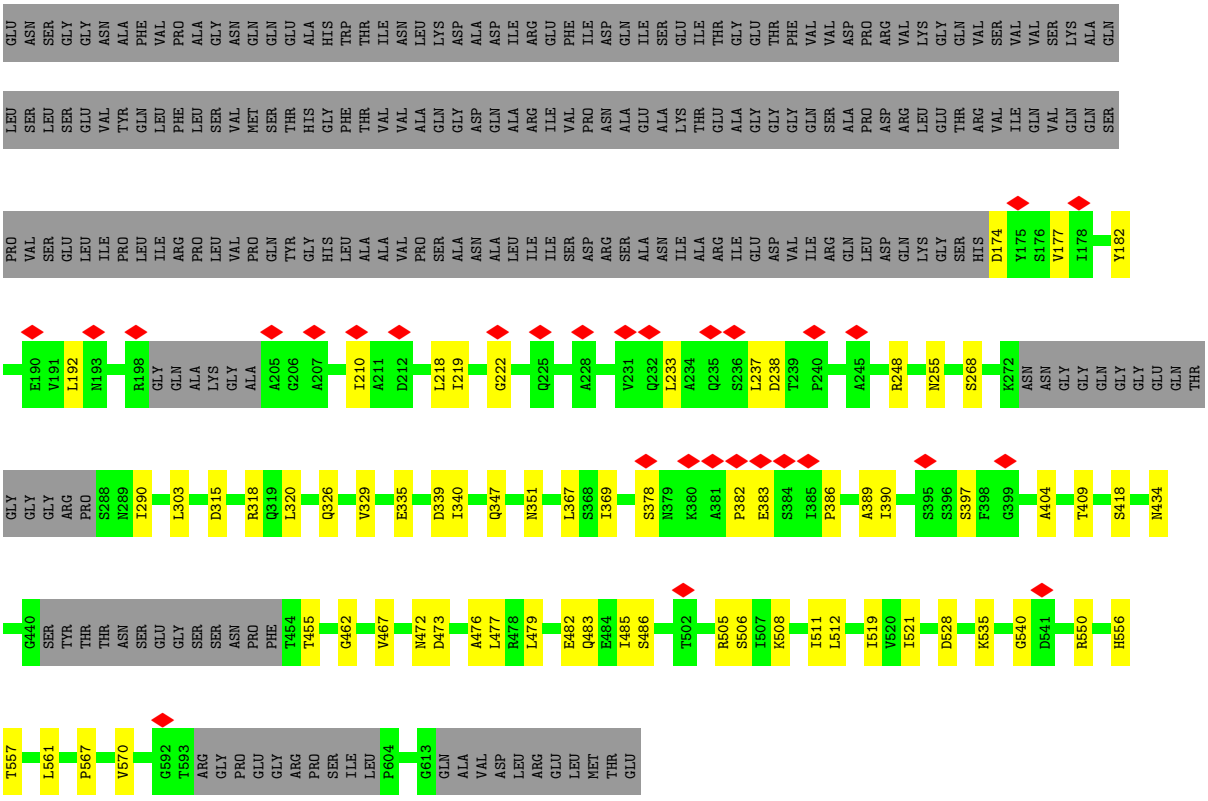
• Molecule 1: Type II secretion system protein D

Chain L:



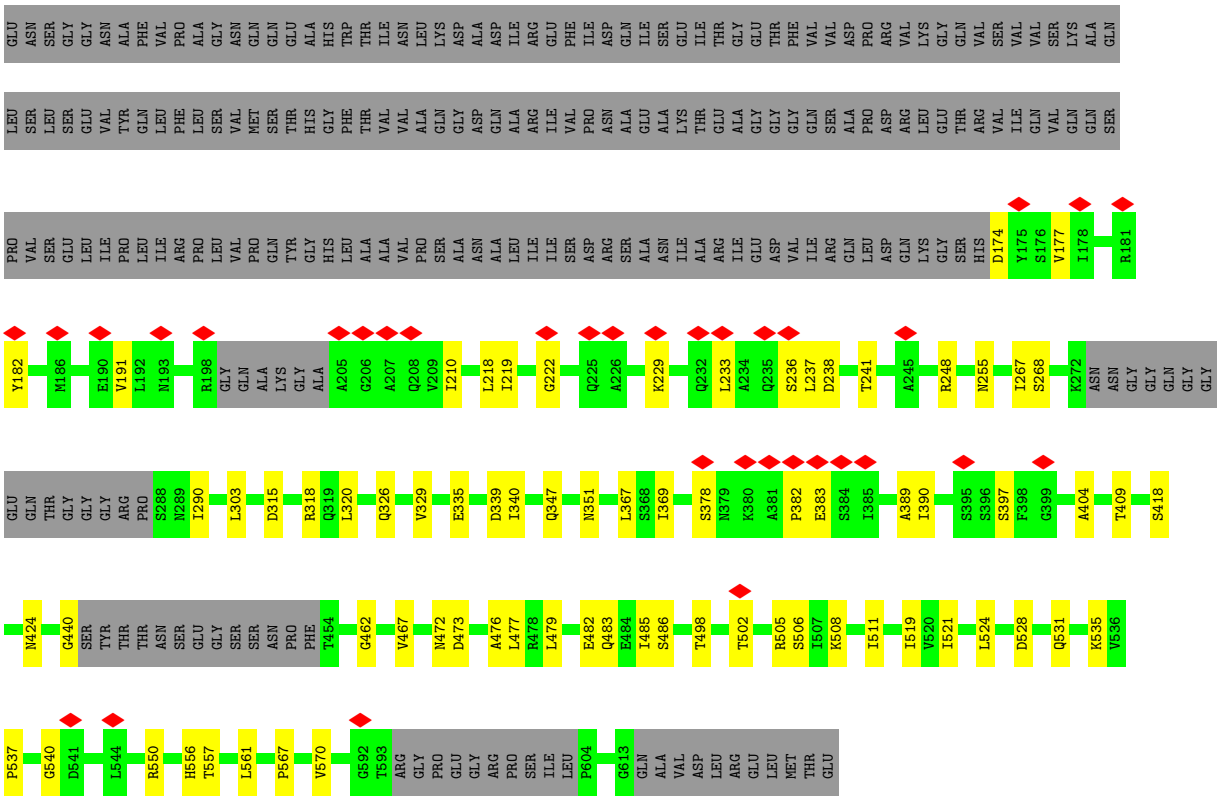
- Molecule 1: Type II secretion system protein D

Chain M:



- Molecule 1: Type II secretion system protein D

Chain N:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C15	Depositor
Number of particles used	18005	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; CTFFIND4	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	600	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	1.501	Depositor
Minimum map value	-0.835	Depositor
Average map value	0.004	Depositor
Map value standard deviation	0.044	Depositor
Recommended contour level	0.176	Depositor
Map size (Å)	343.43997, 343.43997, 343.43997	wwPDB
Map dimensions	324, 324, 324	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.06, 1.06, 1.06	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	0.44	0/2987	0.67	0/4050
1	B	0.44	0/2987	0.67	0/4050
1	C	0.44	0/2987	0.67	0/4050
1	D	0.44	0/2987	0.67	0/4050
1	E	0.44	0/2987	0.67	0/4050
1	F	0.44	0/2987	0.67	0/4050
1	G	0.44	0/2987	0.67	0/4050
1	H	0.44	0/2987	0.67	0/4050
1	I	0.44	0/2987	0.67	0/4050
1	J	0.44	0/2987	0.67	0/4050
1	K	0.44	0/2987	0.67	0/4050
1	L	0.44	0/2987	0.67	0/4050
1	M	0.44	0/2987	0.67	0/4050
1	N	0.44	0/2987	0.67	0/4050
1	O	0.44	0/2987	0.67	0/4050
All	All	0.44	0/44805	0.67	0/60750

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	2961	0	3052	53	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2961	0	3052	60	0
1	C	2961	0	3052	74	0
1	D	2961	0	3052	79	0
1	E	2961	0	3052	63	0
1	F	2961	0	3052	55	0
1	G	2961	0	3052	55	0
1	H	2961	0	3052	53	0
1	I	2961	0	3052	50	0
1	J	2961	0	3052	48	0
1	K	2961	0	3052	48	0
1	L	2961	0	3052	47	0
1	M	2961	0	3052	46	0
1	N	2961	0	3052	56	0
1	O	2961	0	3052	56	0
All	All	44415	0	45780	738	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 738 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:D:512:LEU:HB3	1:E:326:GLN:HE21	1.22	1.04
1:G:440:GLY:HA2	1:H:455:THR:HA	1.51	0.90
1:D:233:LEU:O	1:D:237:LEU:HB2	1.77	0.85
1:F:233:LEU:O	1:F:237:LEU:HB2	1.77	0.85
1:I:233:LEU:O	1:I:237:LEU:HB2	1.77	0.85

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	386/624 (62%)	344 (89%)	42 (11%)	0	100	100
1	B	386/624 (62%)	344 (89%)	42 (11%)	0	100	100
1	C	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	D	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	E	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	F	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	G	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	H	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	I	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	J	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	K	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	L	386/624 (62%)	344 (89%)	42 (11%)	0	100	100
1	M	386/624 (62%)	344 (89%)	42 (11%)	0	100	100
1	N	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
1	O	386/624 (62%)	345 (89%)	41 (11%)	0	100	100
All	All	5790/9360 (62%)	5171 (89%)	619 (11%)	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	326/510 (64%)	326 (100%)	0	100	100
1	B	326/510 (64%)	326 (100%)	0	100	100
1	C	326/510 (64%)	326 (100%)	0	100	100
1	D	326/510 (64%)	326 (100%)	0	100	100
1	E	326/510 (64%)	326 (100%)	0	100	100
1	F	326/510 (64%)	326 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	326/510 (64%)	326 (100%)	0	100	100
1	H	326/510 (64%)	326 (100%)	0	100	100
1	I	326/510 (64%)	326 (100%)	0	100	100
1	J	326/510 (64%)	326 (100%)	0	100	100
1	K	326/510 (64%)	326 (100%)	0	100	100
1	L	326/510 (64%)	326 (100%)	0	100	100
1	M	326/510 (64%)	326 (100%)	0	100	100
1	N	326/510 (64%)	326 (100%)	0	100	100
1	O	326/510 (64%)	326 (100%)	0	100	100
All	All	4890/7650 (64%)	4890 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 57 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	472	ASN
1	N	483	GLN
1	H	608	ASN
1	N	472	ASN
1	M	472	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.

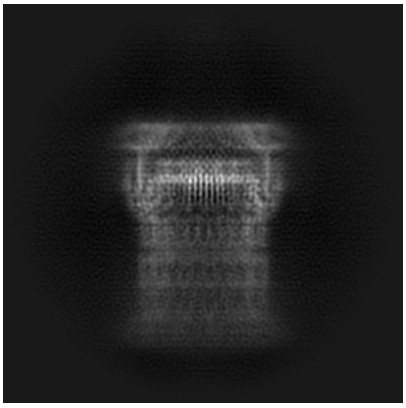
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-8860. These allow visual inspection of the internal detail of the map and identification of artifacts.

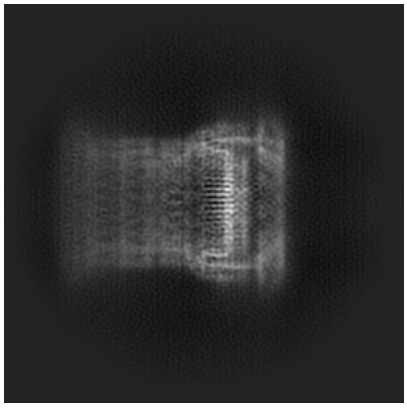
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

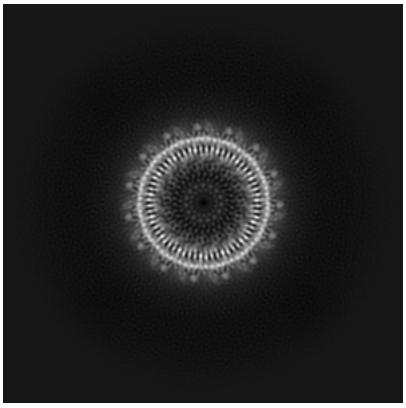
6.1.1 Primary map



X



Y

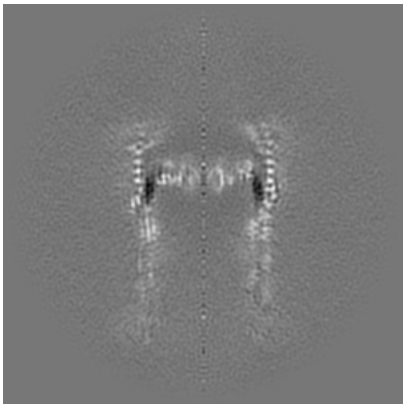


Z

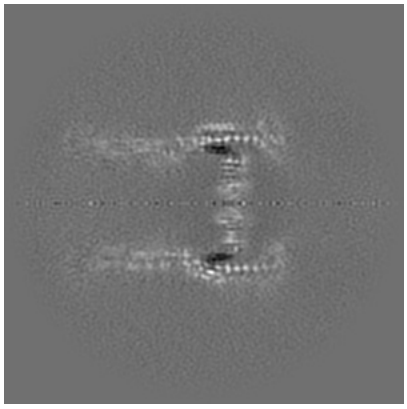
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

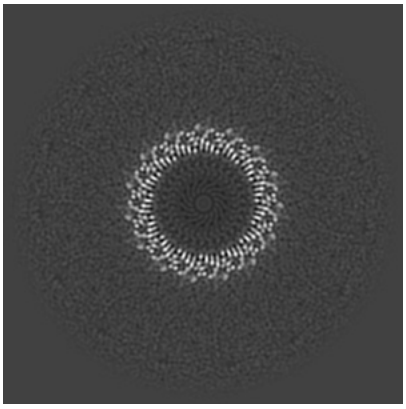
6.2.1 Primary map



X Index: 162



Y Index: 162

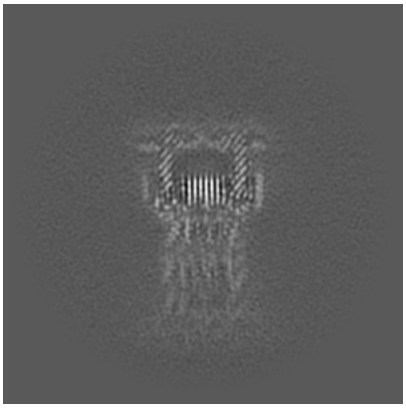


Z Index: 162

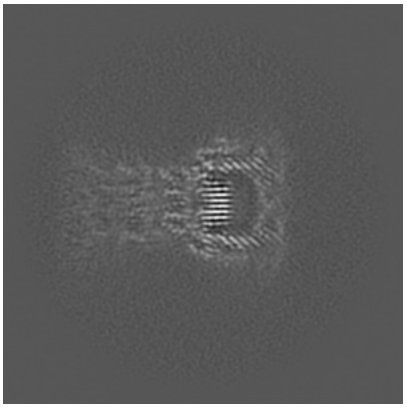
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

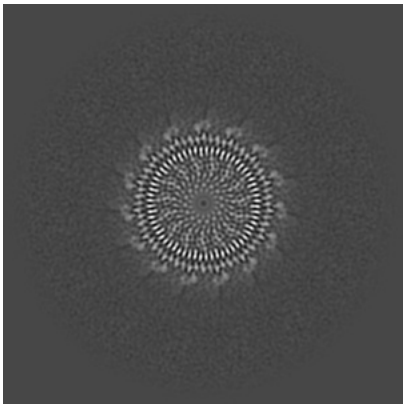
6.3.1 Primary map



X Index: 205



Y Index: 205

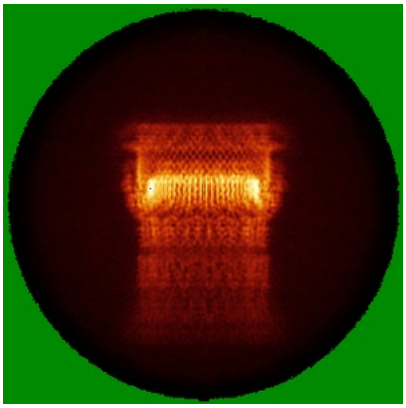


Z Index: 180

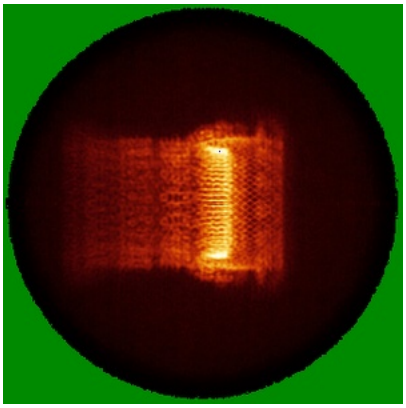
The images above show the largest variance slices of the map in three orthogonal directions.

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

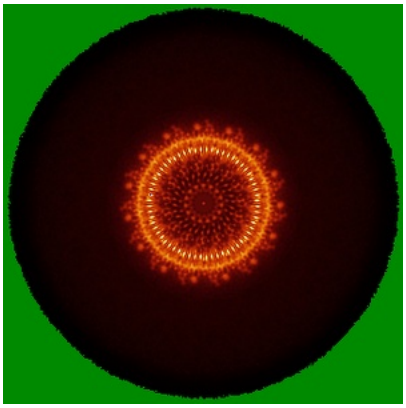
6.4.1 Primary map



X



Y

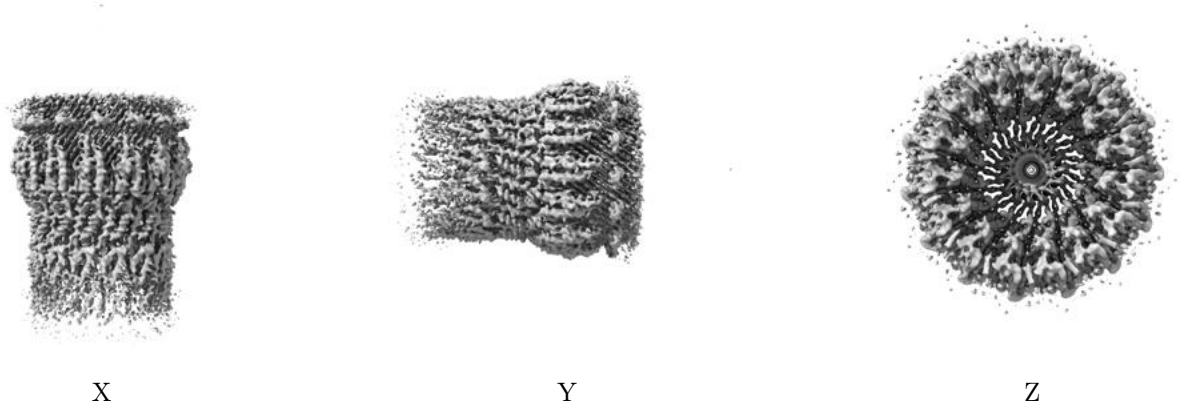


Z

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

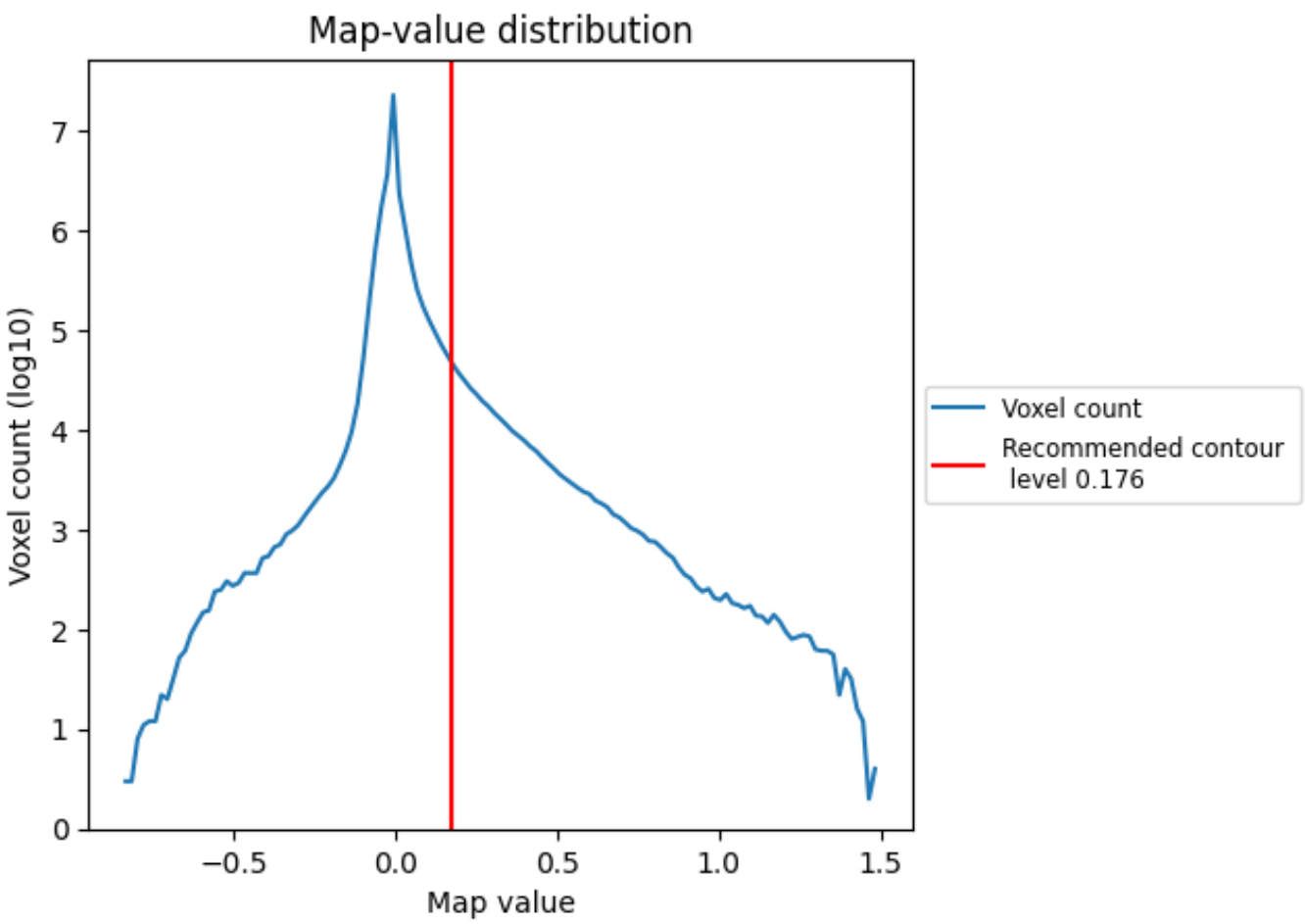
6.6 Mask visualisation (i)

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

7 Map analysis (i)

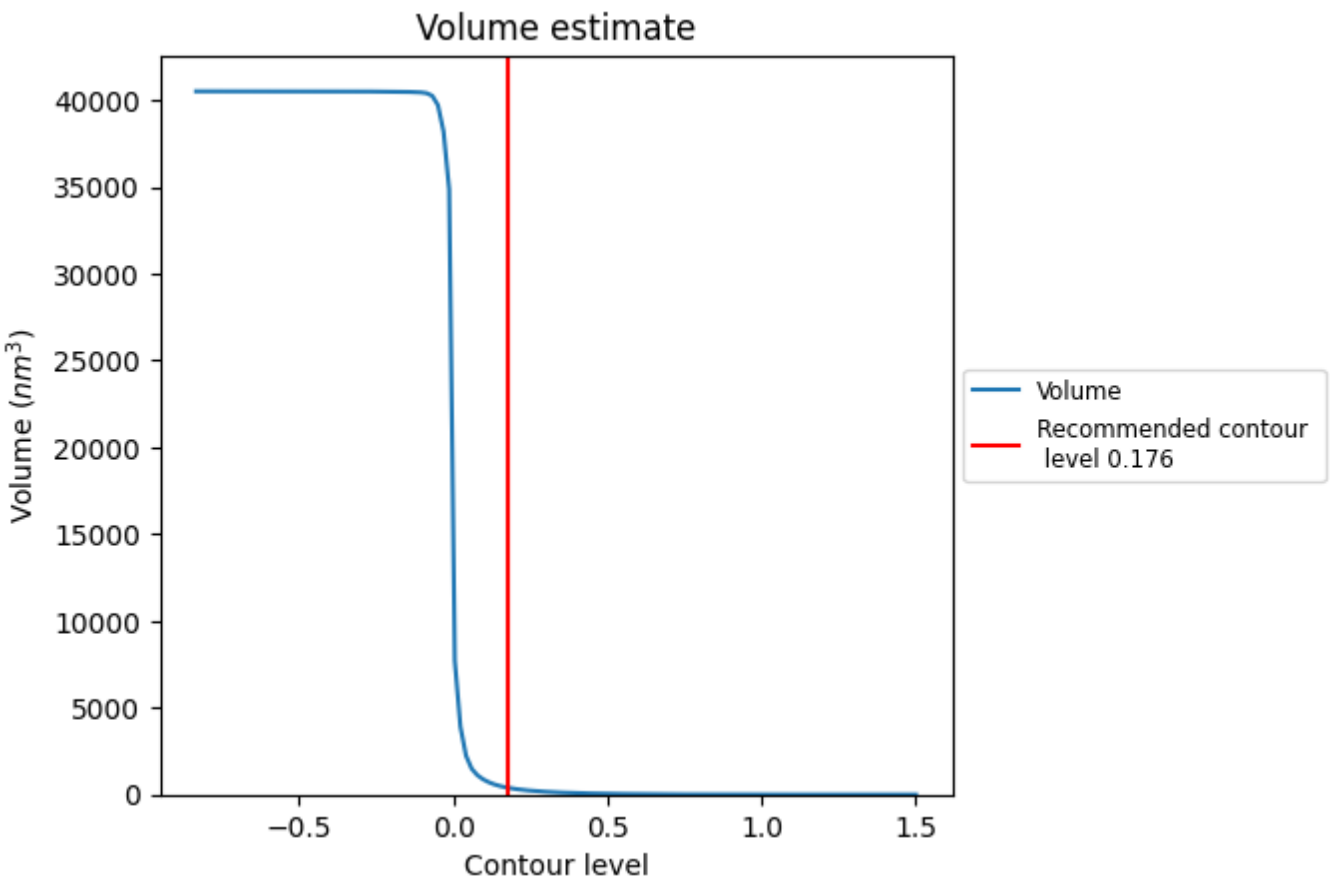
This section contains the results of statistical analysis of the map.

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.

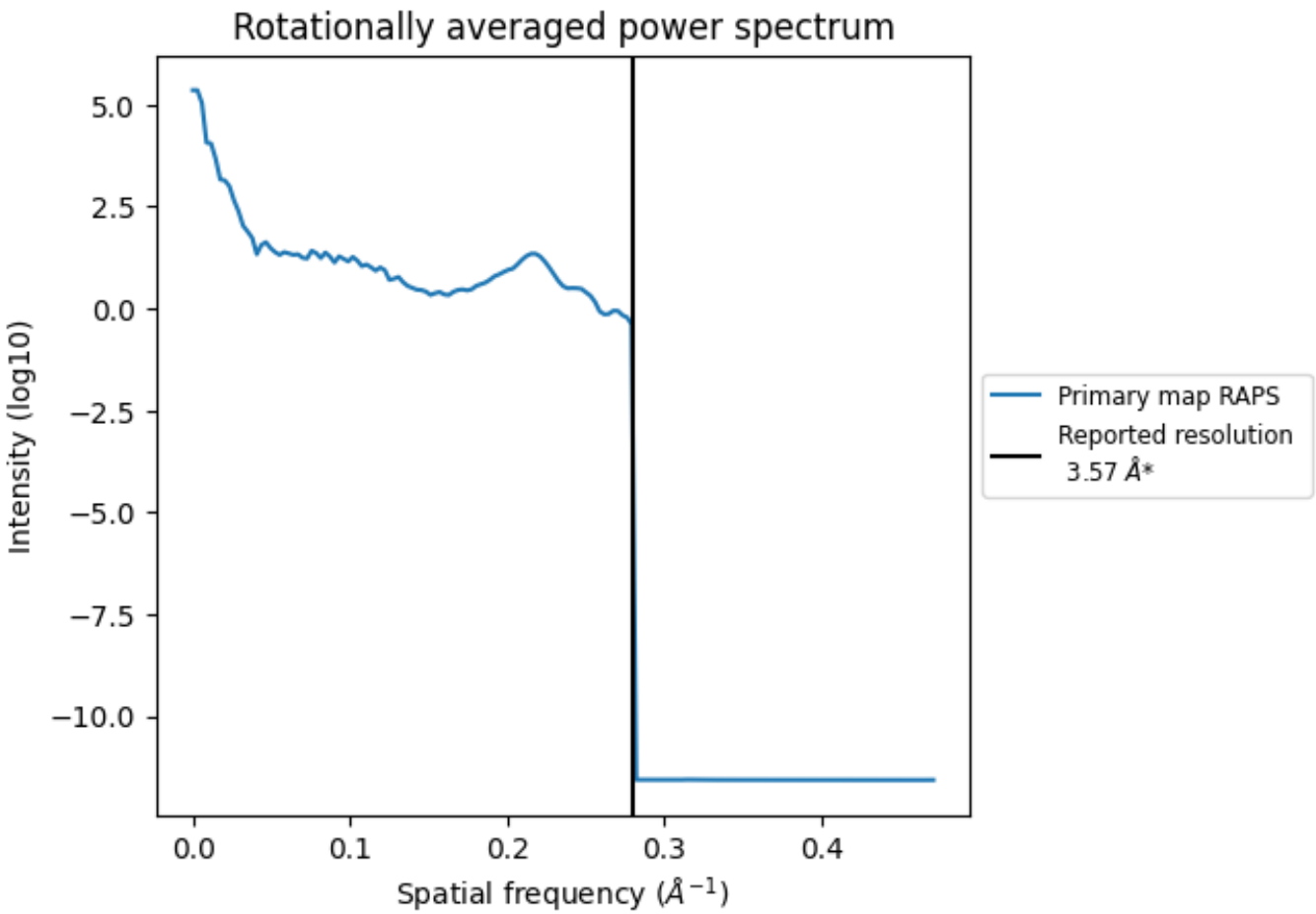
7.2 Volume estimate (i)



The volume at the recommended contour level is 397 nm^3; this corresponds to an approximate mass of 359 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)

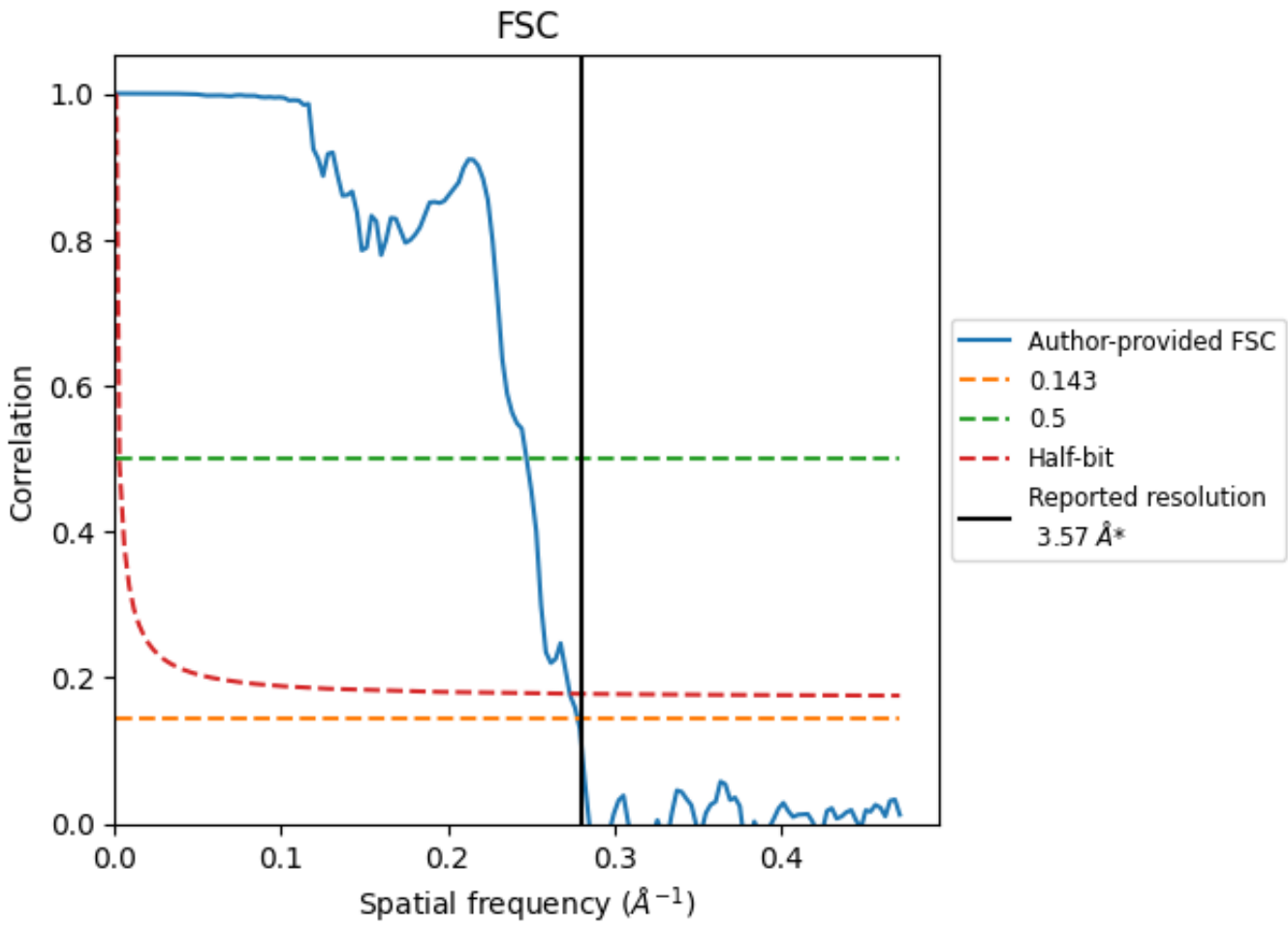


*Reported resolution corresponds to spatial frequency of 0.280 Å^{-1}

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 0.280 Å^{-1}

8.2 Resolution estimates (i)

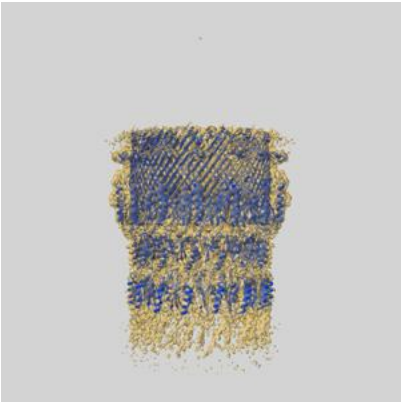
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.57	-	-
Author-provided FSC curve	3.59	4.04	3.66
Unmasked-calculated*	-	-	-

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

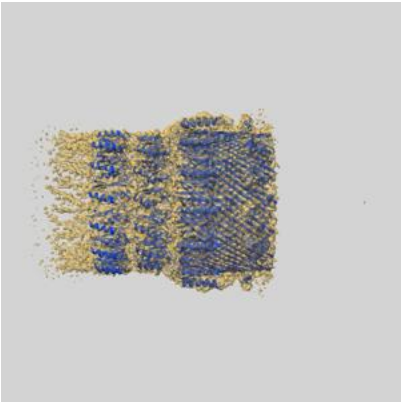
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-8860 and PDB model 5WLN. Per-residue inclusion information can be found in section 3 on page 5.

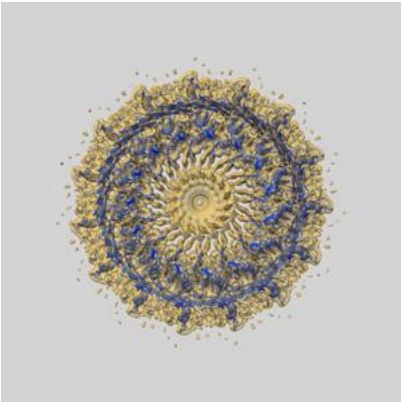
9.1 Map-model overlay (i)



X



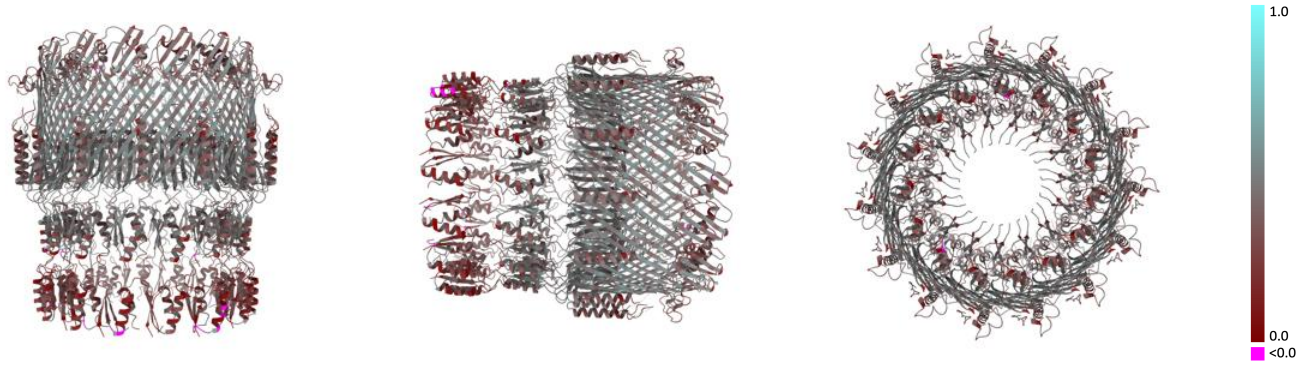
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.176 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.

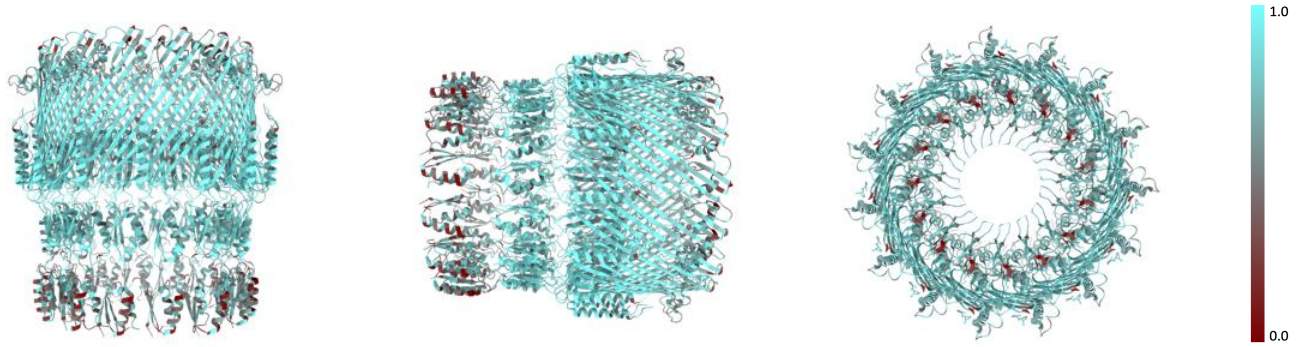
9.2 Q-score mapped to coordinate model (i)



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.

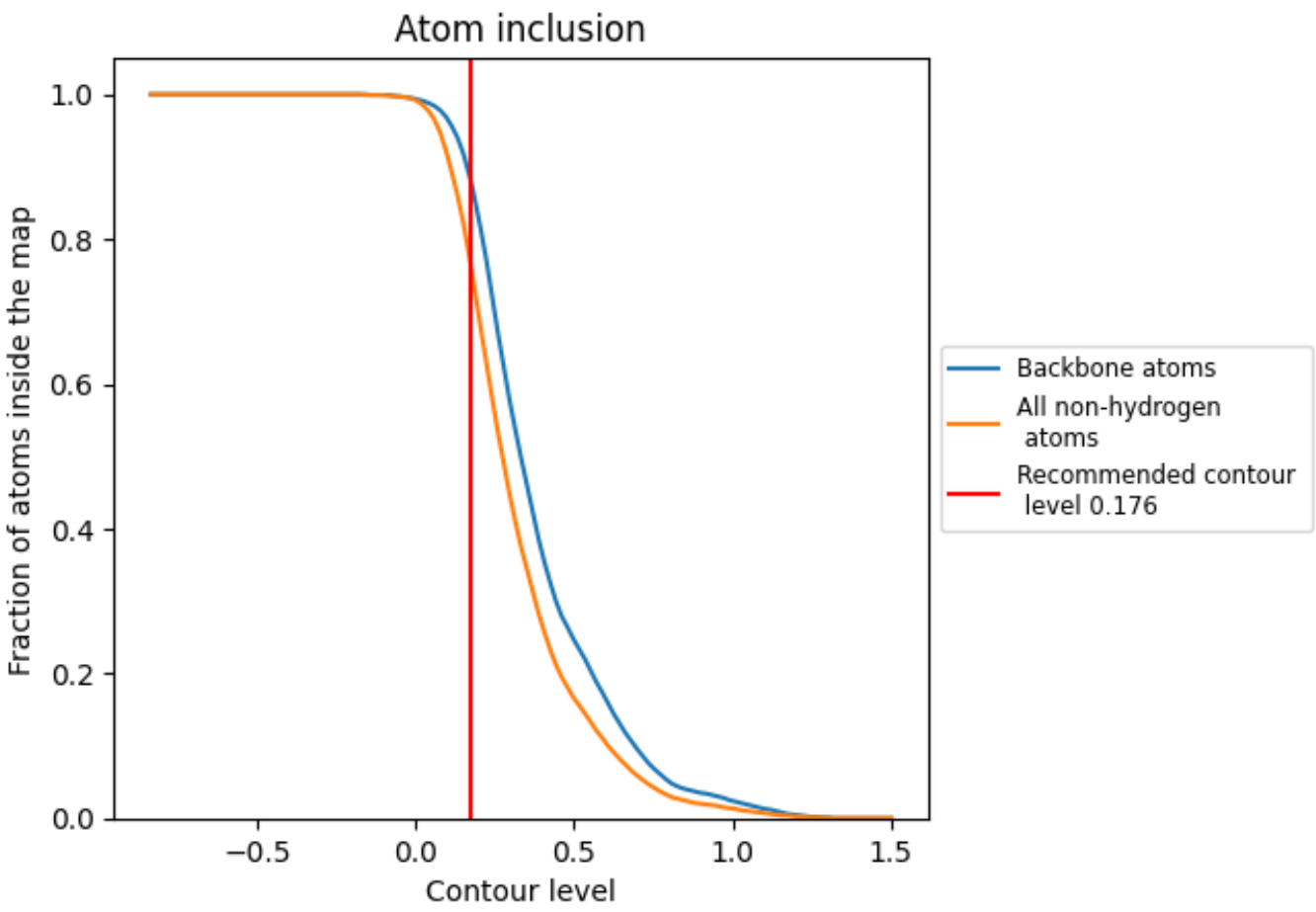
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.176).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 76% 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 (0.176) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7600	0.3960
A	0.7580	0.3930
B	0.7620	0.3980
C	0.7510	0.3910
D	0.7600	0.3930
E	0.7390	0.3660
F	0.7620	0.3990
G	0.7700	0.4050
H	0.7650	0.4010
I	0.7660	0.4010
J	0.7670	0.4010
K	0.7670	0.4030
L	0.7610	0.4060
M	0.7630	0.4030
N	0.7550	0.3850
O	0.7550	0.3910