wwPDB EM Validation Summary Report

Oct 27, 2024 – 03:18 PM JST

PDB ID	:	7CUN
EMDB ID	:	EMD-30473
Title	:	The structure of human Integrator-PP2A complex
Authors	:	Zheng, H.; Qi, Y.; Liu, W.; Li, J.; Wang, J.; Xu, Y.
Deposited on	:	2020-08-23
Resolution	:	3.50 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:

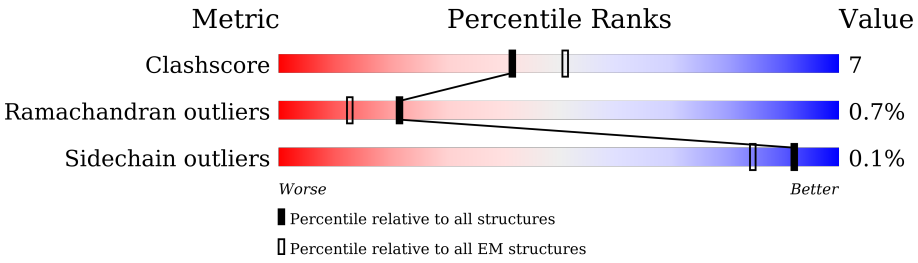
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.50 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.



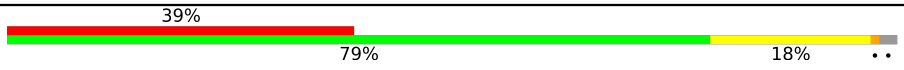
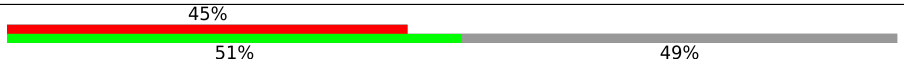
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	2190	29% 67%
2	B	1204	75% 13% 12%
3	D	963	5% 69% 16% 14%
4	E	1019	67% 10% 22%
5	F	887	50% 10% 40%
6	G	962	79% 14% 7%
7	H	995	78% 13% 8%
8	I	658	24% 68% 26% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	K	600	
10	P	589	
11	Q	309	
12	U	594	

2 Entry composition (i)

There are 14 unique types of molecules in this entry. The entry contains 57642 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 Integrator complex subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	730	4653	2872	876	880	25	0	0

- Molecule 2 is a protein called Integrator complex subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1058	7541	4765	1318	1407	51	0	0

- Molecule 3 is a protein called Integrator complex subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	827	5942	3743	1049	1119	31	0	0

- Molecule 4 is a protein called Integrator complex subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	796	5288	3321	989	963	15	0	0

- Molecule 5 is a protein called Integrator complex subunit 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	534	3822	2417	669	715	21	0	0

- Molecule 6 is a protein called Integrator complex subunit 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	G	895	6826	4319	1196	1270	41	0	0

- Molecule 7 is a protein called Integrator complex subunit 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	912	Total	C	N	O	S	0	0
			6657	4216	1178	1231	32		

- Molecule 8 is a protein called Integrator complex subunit 9.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	I	624	Total	C	N	O	S	0	0
			4701	3004	778	888	31		

- Molecule 9 is a protein called Integrator complex subunit 11.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	K	588	Total	C	N	O	S	0	0
			3792	2347	709	721	15		

- Molecule 10 is a protein called PP2A-A.

Mol	Chain	Residues	Atoms						AltConf	Trace
10	P	582	Total	C	N	O	S	Se	0	0
			4535	2881	764	863	14	13		

- Molecule 11 is a protein called PP2A-C.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	Q	293	Total	C	N	O	S	0	0
			2366	1497	405	449	15		

- Molecule 12 is a protein called unknown.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	U	303	Total	C	N	O	0	0
			1515	909	303	303		

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
13	K	2	Total	Zn	0
			2	2	

- Molecule 14 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

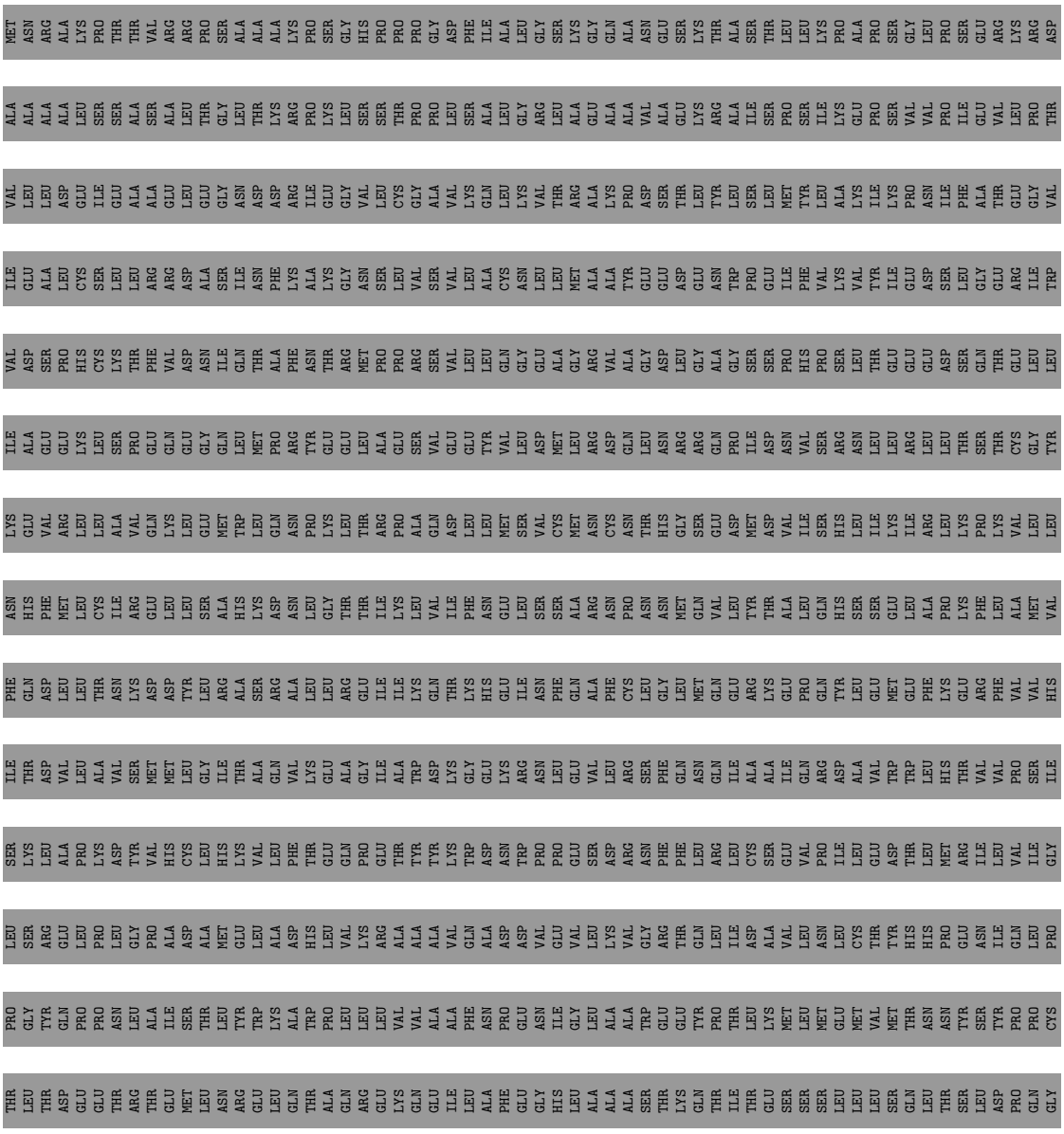
Mol	Chain	Residues	Atoms		AltConf
14	Q	2	Total	Mn	0
			2	2	

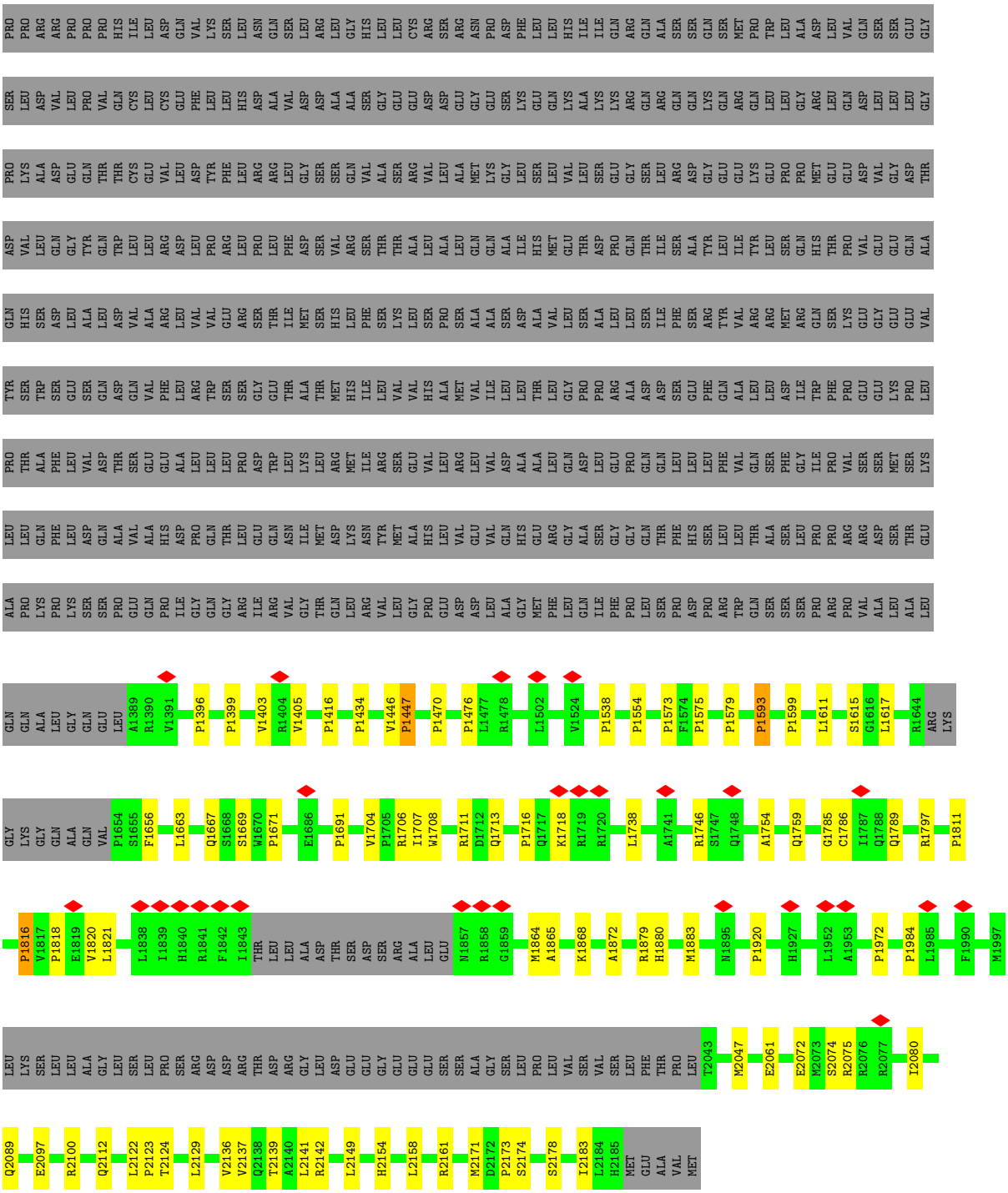
3 Residue-property plots (i)

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: Integrator complex subunit 1

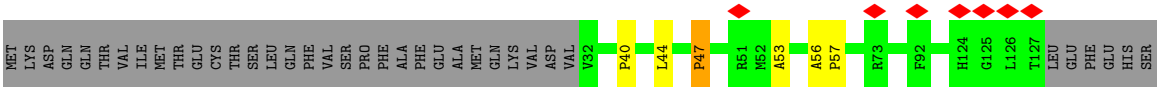
Chain A:

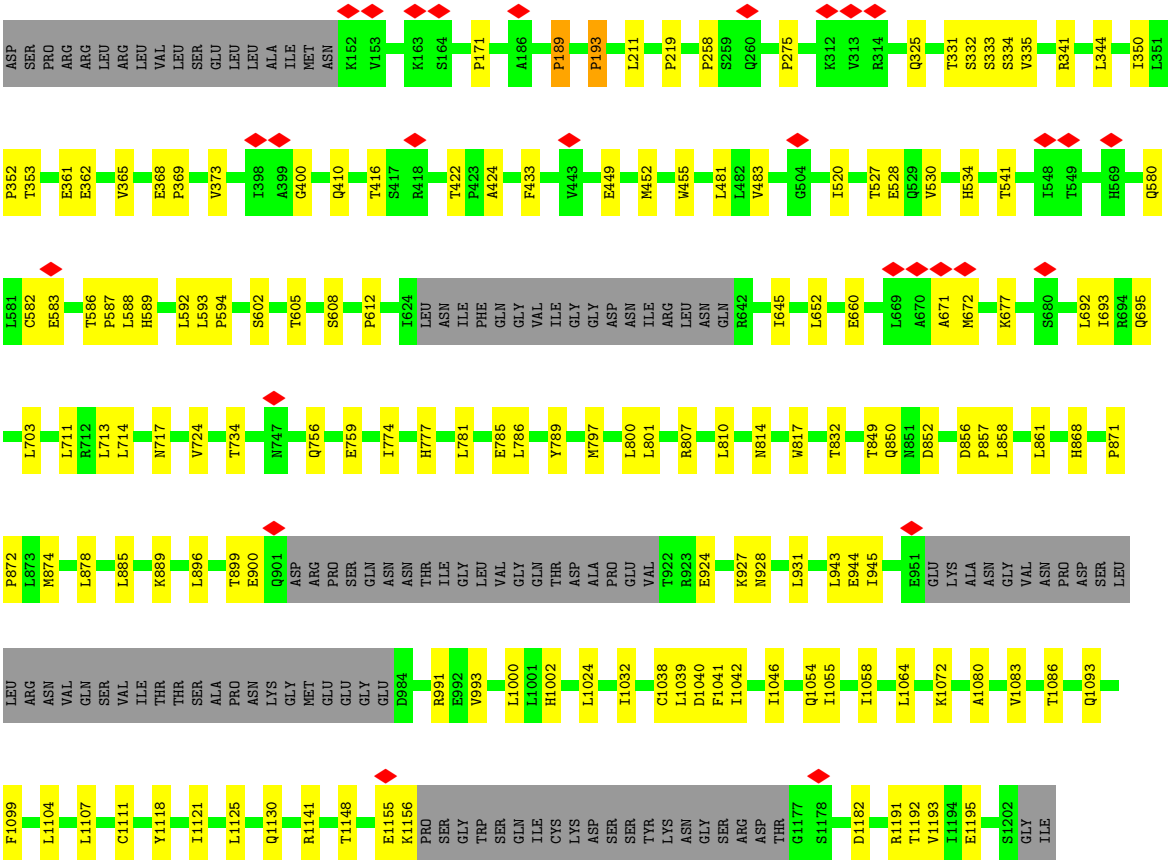




• Molecule 2: Integrator complex subunit 2

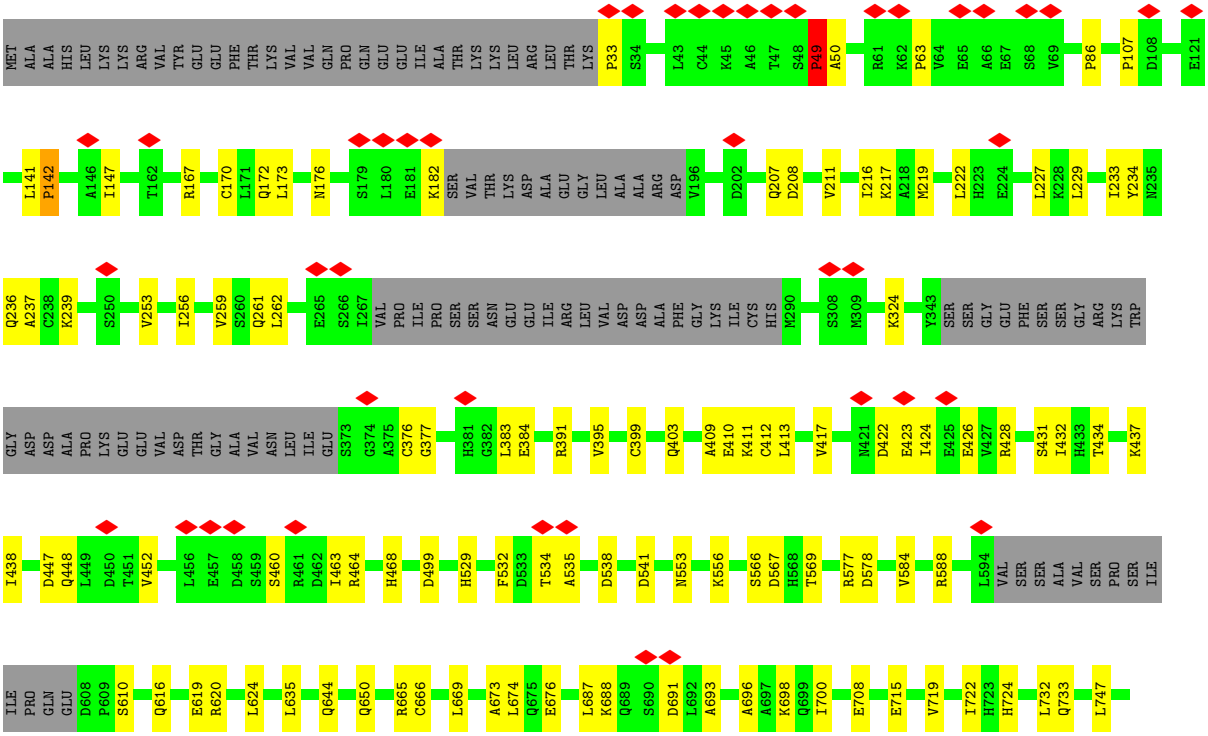
Chain B:

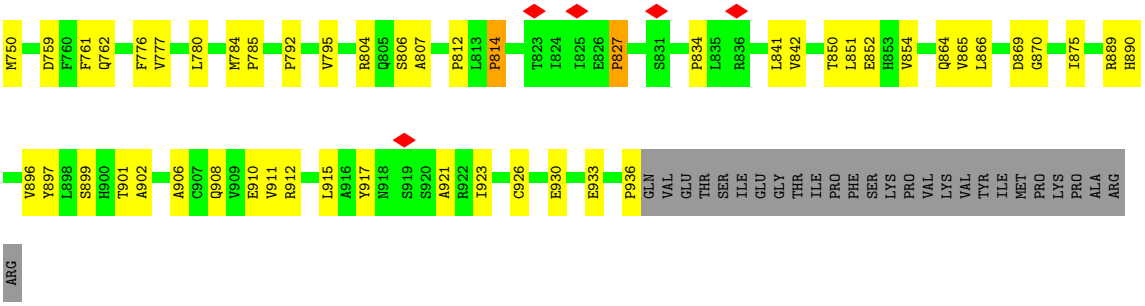




• Molecule 3: Integrator complex subunit 4

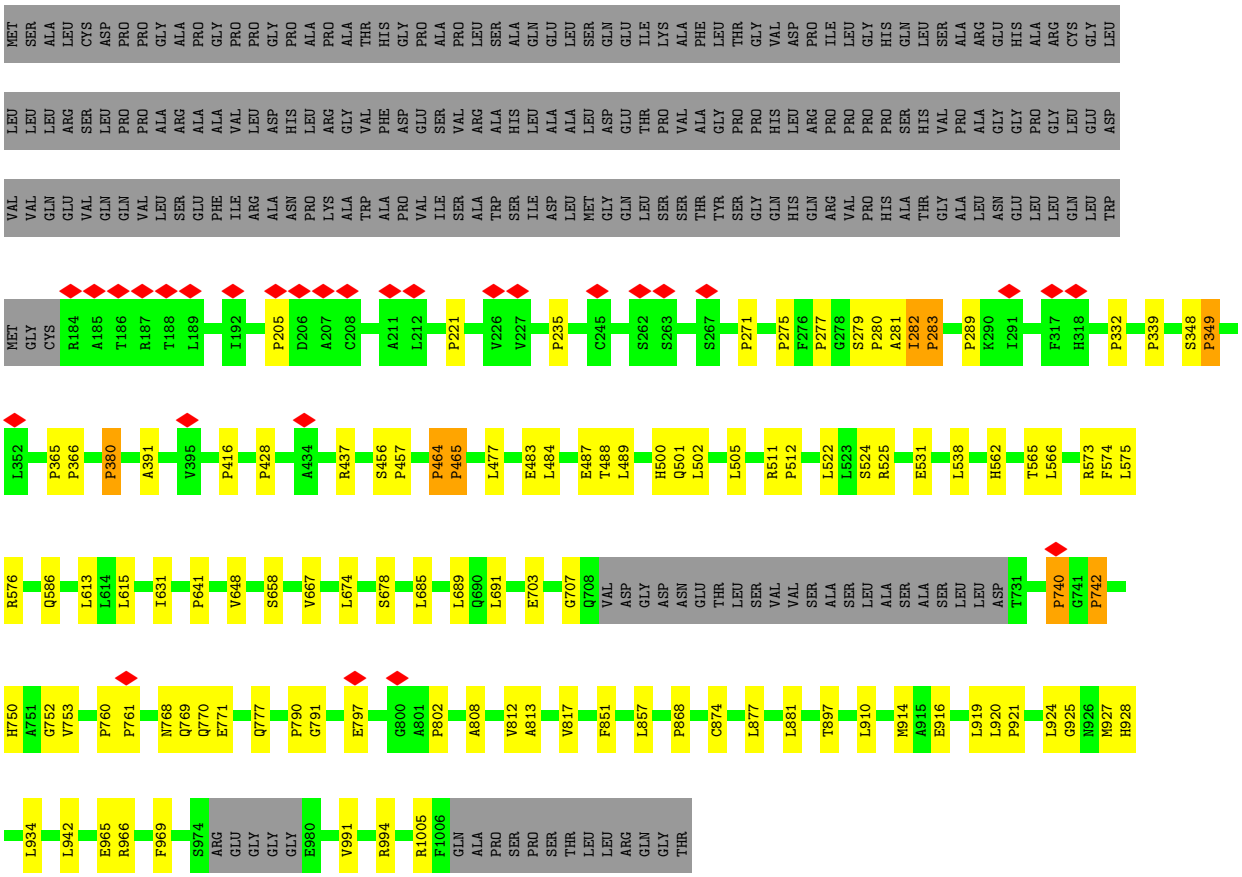
Chain D:





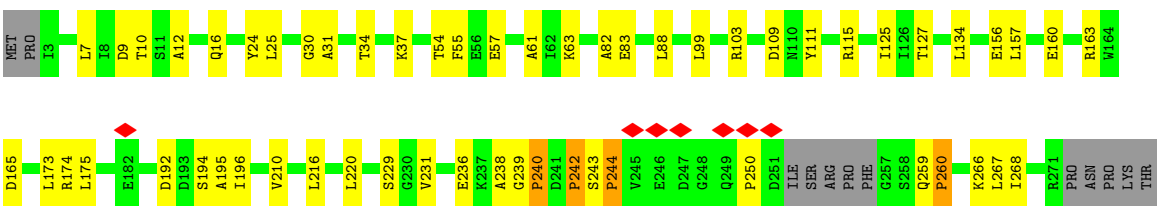
• Molecule 4: Integrator complex subunit 5

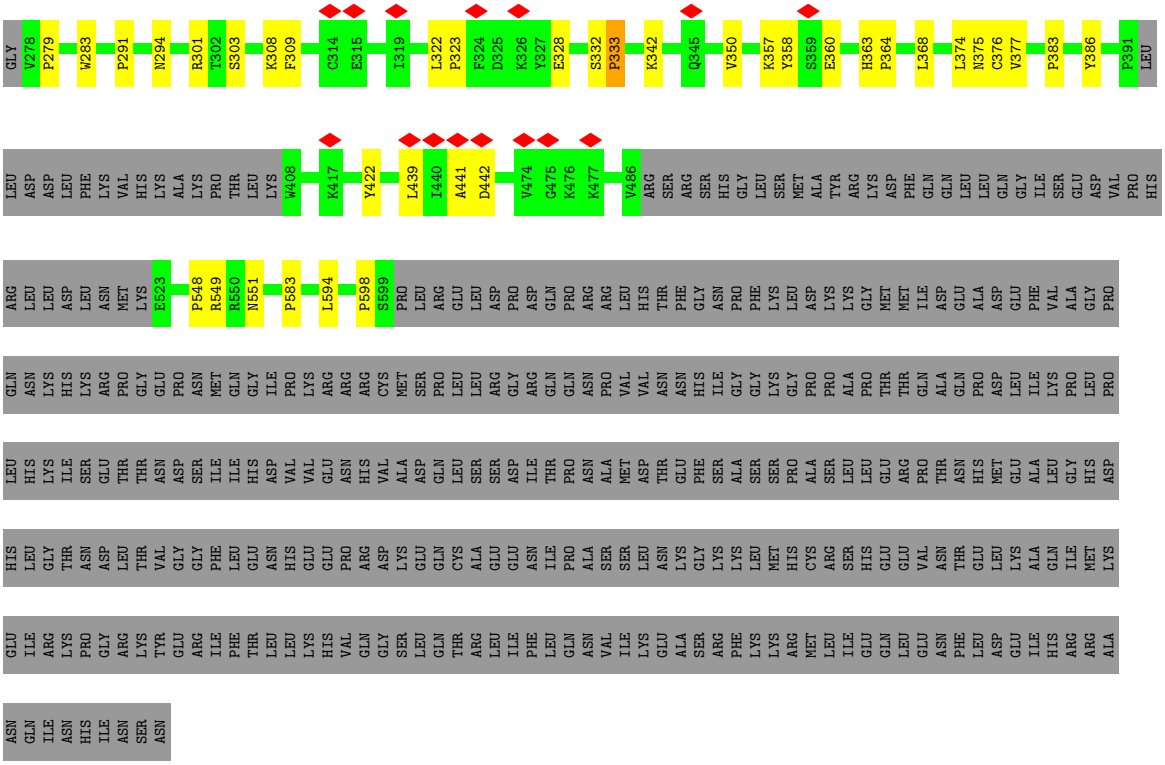
Chain E:



• Molecule 5: Integrator complex subunit 6

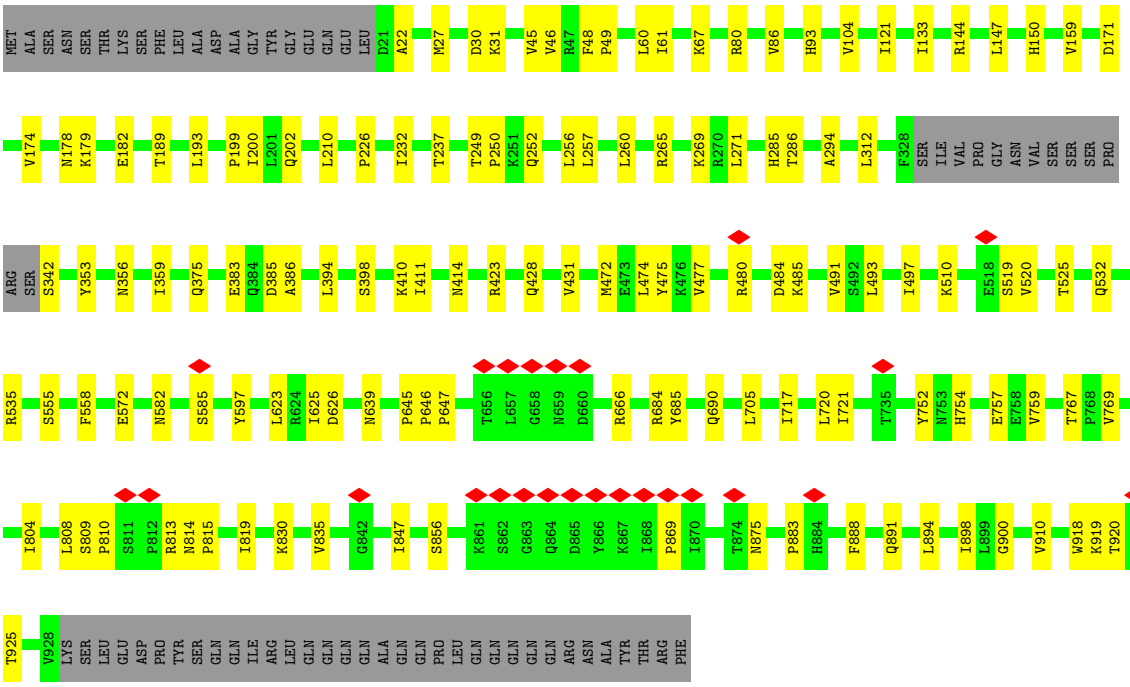
Chain F:





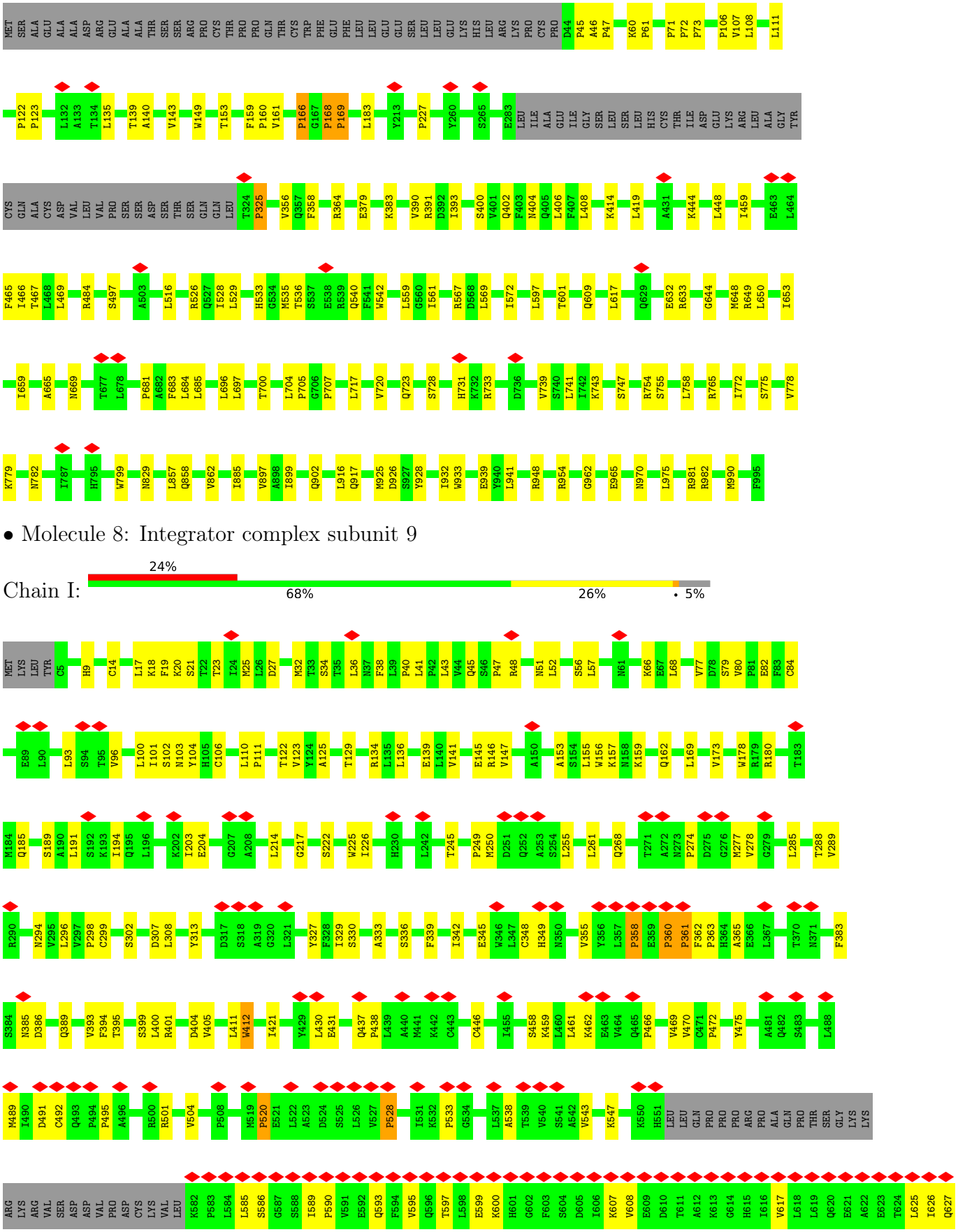
• Molecule 6: Integrator complex subunit 7

Chain G:



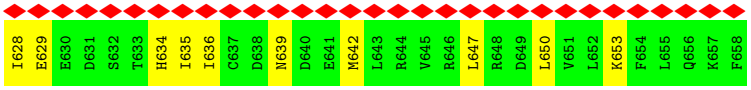
• Molecule 7: Integrator complex subunit 8

Chain H:



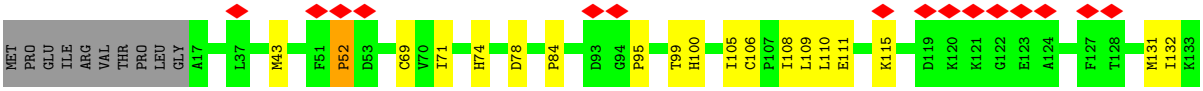
• Molecule 8: Integrator complex subunit 9

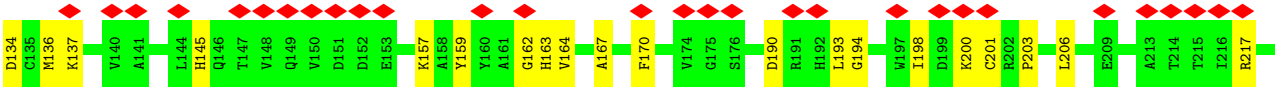
Chain I:

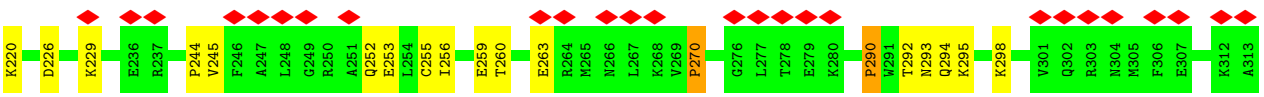


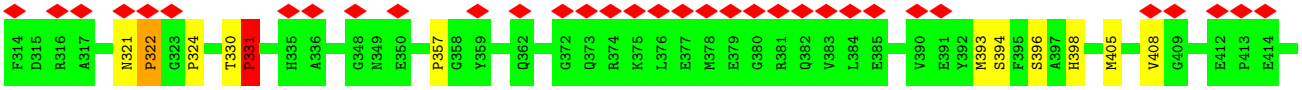
• Molecule 9: Integrator complex subunit 11

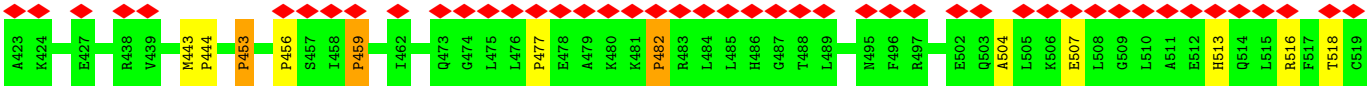
Chain K:

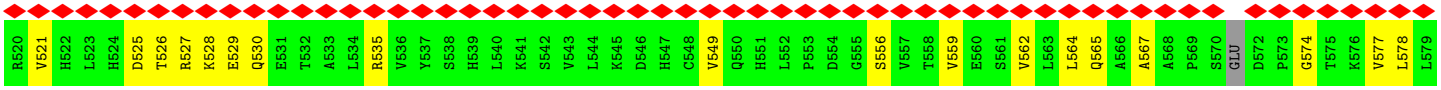


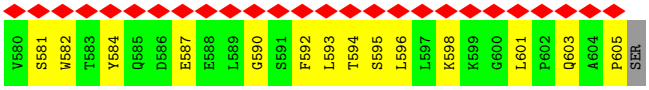






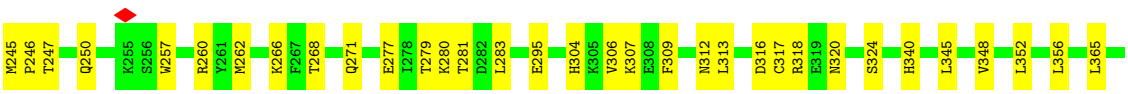


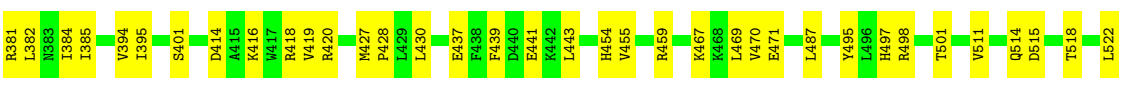


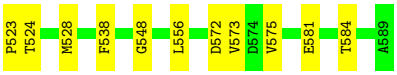


• Molecule 10: PP2A-A

Chain P:

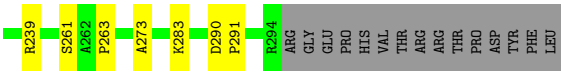






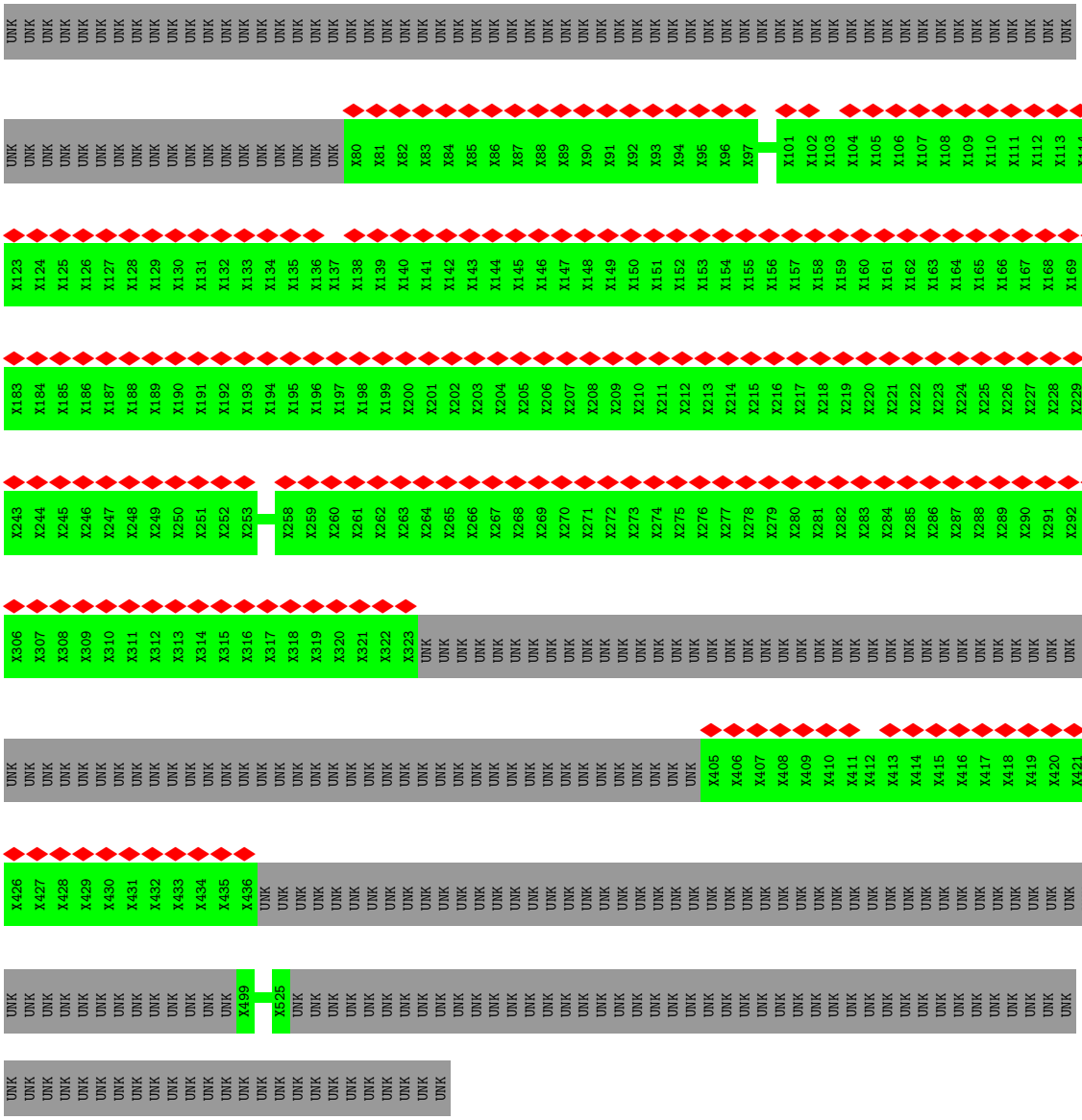
• Molecule 11: PP2A-C

Chain Q:



• Molecule 12: unknown

Chain U:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	68378	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50.0	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	0.209	Depositor
Minimum map value	-0.117	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.016	Depositor
Map size (Å)	506.39996, 506.39996, 506.39996	wwPDB
Map dimensions	480, 480, 480	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.055, 1.055, 1.055	Depositor

5 Model quality

5.1 Standard geometry

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

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	0.24	0/4706	0.51	20/6443 (0.3%)
2	B	0.24	0/7651	0.45	9/10439 (0.1%)
3	D	0.24	0/6039	0.47	10/8229 (0.1%)
4	E	0.25	0/5387	0.50	22/7365 (0.3%)
5	F	0.25	0/3903	0.51	9/5328 (0.2%)
6	G	0.23	0/6937	0.40	4/9397 (0.0%)
7	H	0.24	0/6765	0.45	13/9207 (0.1%)
8	I	0.25	0/4807	0.54	7/6562 (0.1%)
9	K	0.25	0/3843	0.59	16/5249 (0.3%)
10	P	0.23	0/4596	0.41	0/6218
11	Q	0.24	0/2423	0.40	0/3285
All	All	0.24	0/57057	0.48	110/77722 (0.1%)

There are no bond length outliers.

The worst 5 of 110 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	K	331	PRO	N-CA-CB	6.96	111.66	103.30
9	K	482	PRO	N-CA-CB	6.92	111.60	103.30
2	B	189	PRO	N-CA-CB	6.89	111.57	103.30
7	H	325	PRO	N-CA-CB	6.88	111.55	103.30
8	I	360	PRO	N-CA-CB	6.87	111.55	103.30

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	4653	0	3696	41	0
2	B	7541	0	7031	90	0
3	D	5942	0	5402	98	0
4	E	5288	0	4612	51	0
5	F	3822	0	3415	49	0
6	G	6826	0	6830	78	0
7	H	6657	0	6168	80	0
8	I	4701	0	4503	112	0
9	K	3792	0	3006	64	0
10	P	4535	0	4637	87	0
11	Q	2366	0	2269	22	0
12	U	1515	0	322	0	0
13	K	2	0	0	0	0
14	Q	2	0	0	0	0
All	All	57642	0	51891	723	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 7.

The worst 5 of 723 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)
11:Q:107:TYR:HB3	11:Q:110:ARG:HB2	1.67	0.75
8:I:41:LEU:HD22	8:I:84:CYS:HB2	1.72	0.72
8:I:101:ILE:HD13	8:I:125:ALA:HB2	1.71	0.71
9:K:528:LYS:HD3	9:K:530:GLN:H	1.54	0.71
1:A:2141:LEU:HD11	1:A:2173:PRO:HB3	1.72	0.70

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	722/2190 (33%)	678 (94%)	38 (5%)	6 (1%)	16 51
2	B	1046/1204 (87%)	965 (92%)	78 (8%)	3 (0%)	37 68
3	D	817/963 (85%)	752 (92%)	59 (7%)	6 (1%)	19 53
4	E	790/1019 (78%)	742 (94%)	38 (5%)	10 (1%)	10 41
5	F	524/887 (59%)	466 (89%)	50 (10%)	8 (2%)	8 39
6	G	891/962 (93%)	845 (95%)	46 (5%)	0	100 100
7	H	908/995 (91%)	852 (94%)	50 (6%)	6 (1%)	19 53
8	I	620/658 (94%)	527 (85%)	86 (14%)	7 (1%)	12 45
9	K	584/600 (97%)	522 (89%)	51 (9%)	11 (2%)	6 34
10	P	580/589 (98%)	567 (98%)	13 (2%)	0	100 100
11	Q	291/309 (94%)	280 (96%)	11 (4%)	0	100 100
All	All	7773/10376 (75%)	7196 (93%)	520 (7%)	57 (1%)	21 53

5 of 57 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1593	PRO
1	A	1816	PRO
2	B	47	PRO
2	B	189	PRO
3	D	49	PRO

5.3.2 Protein sidechains

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	313/1907 (16%)	313 (100%)	0	100 100
2	B	711/1072 (66%)	711 (100%)	0	100 100
3	D	543/845 (64%)	542 (100%)	1 (0%)	92 97
4	E	410/812 (50%)	410 (100%)	0	100 100
5	F	353/796 (44%)	353 (100%)	0	100 100
6	G	734/840 (87%)	734 (100%)	0	100 100
7	H	624/896 (70%)	624 (100%)	0	100 100
8	I	505/600 (84%)	502 (99%)	3 (1%)	84 91
9	K	258/520 (50%)	258 (100%)	0	100 100
10	P	509/498 (102%)	509 (100%)	0	100 100
11	Q	259/274 (94%)	259 (100%)	0	100 100
All	All	5219/9060 (58%)	5215 (100%)	4 (0%)	92 97

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	182	LYS
8	I	412	TRP
8	I	459	LYS
8	I	607	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 94 such sidechains are listed below:

Mol	Chain	Res	Type
7	H	404	ASN
8	I	468	HIS
7	H	609	GLN
7	H	902	GLN
9	K	192	HIS

5.3.3 RNA (i)

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 4 ligands modelled in this entry, 4 are 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.

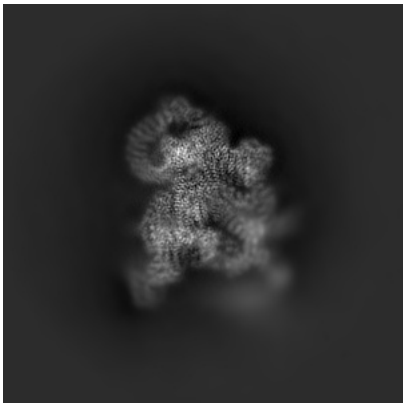
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-30473. These allow visual inspection of the internal detail of the map and identification of artifacts.

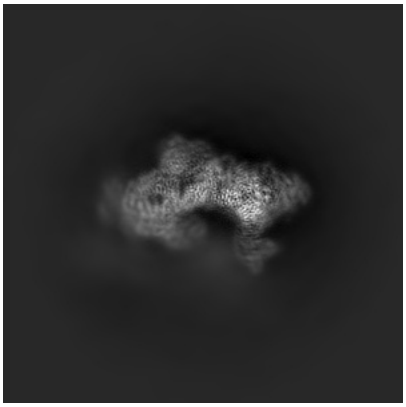
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

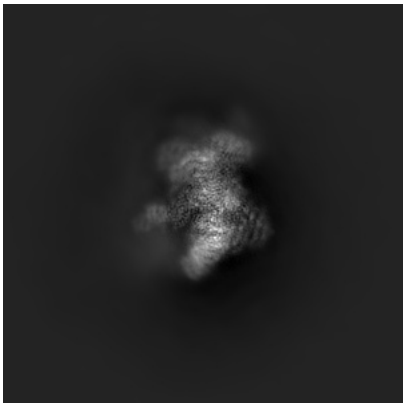
6.1.1 Primary map



X



Y

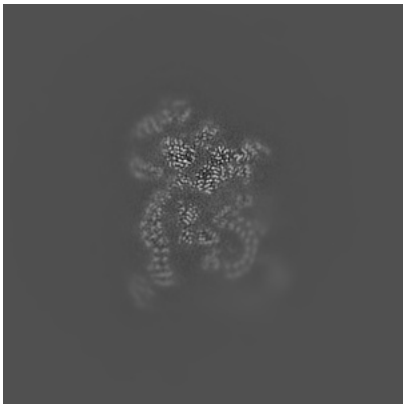


Z

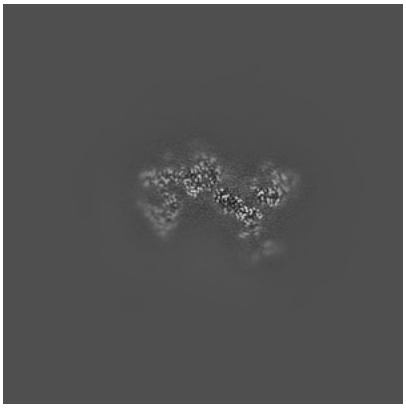
The images above show the map projected in three orthogonal directions.

6.2 Central slices

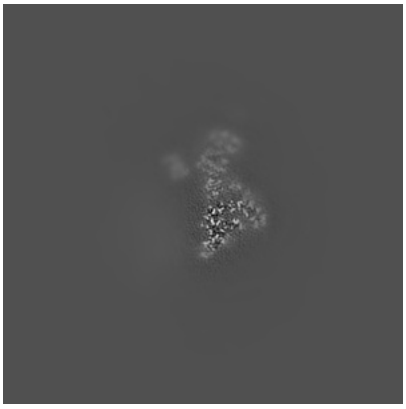
6.2.1 Primary map



X Index: 240



Y Index: 240

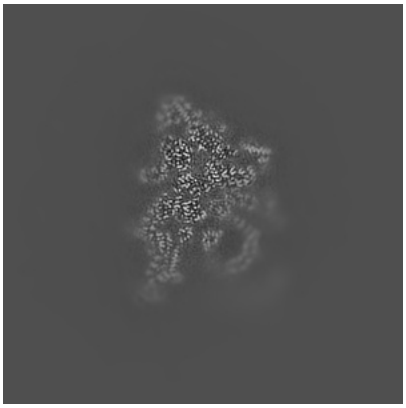


Z Index: 240

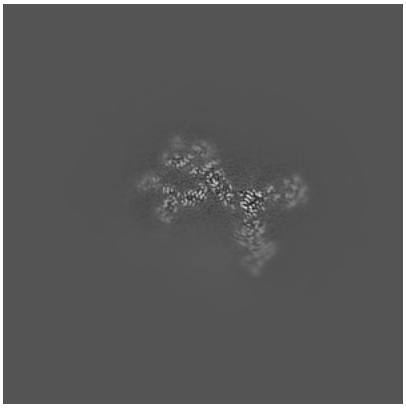
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

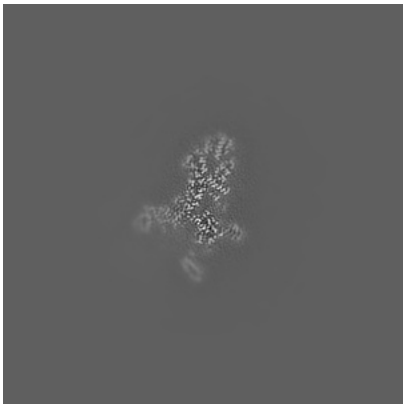
6.3.1 Primary map



X Index: 251



Y Index: 220

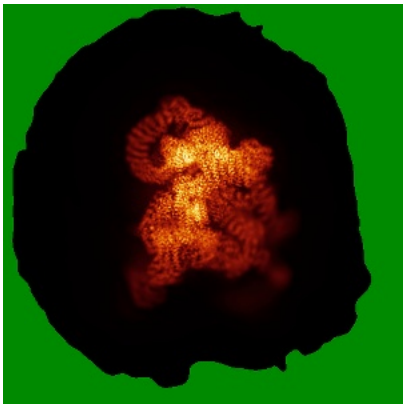


Z Index: 297

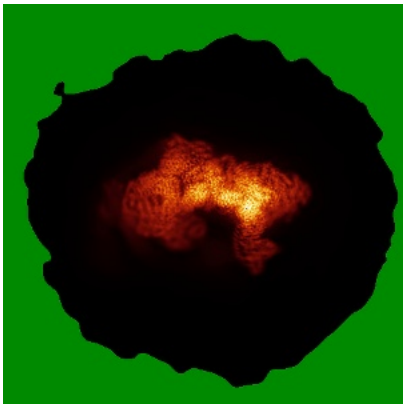
The images above show the largest variance slices of the map in three orthogonal directions.

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

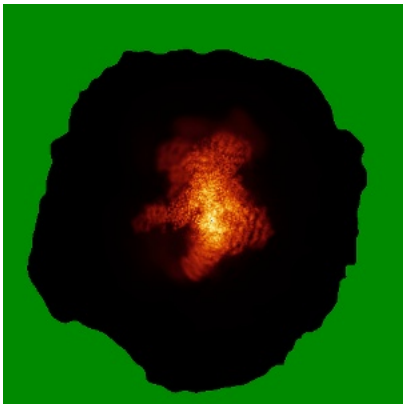
6.4.1 Primary map



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

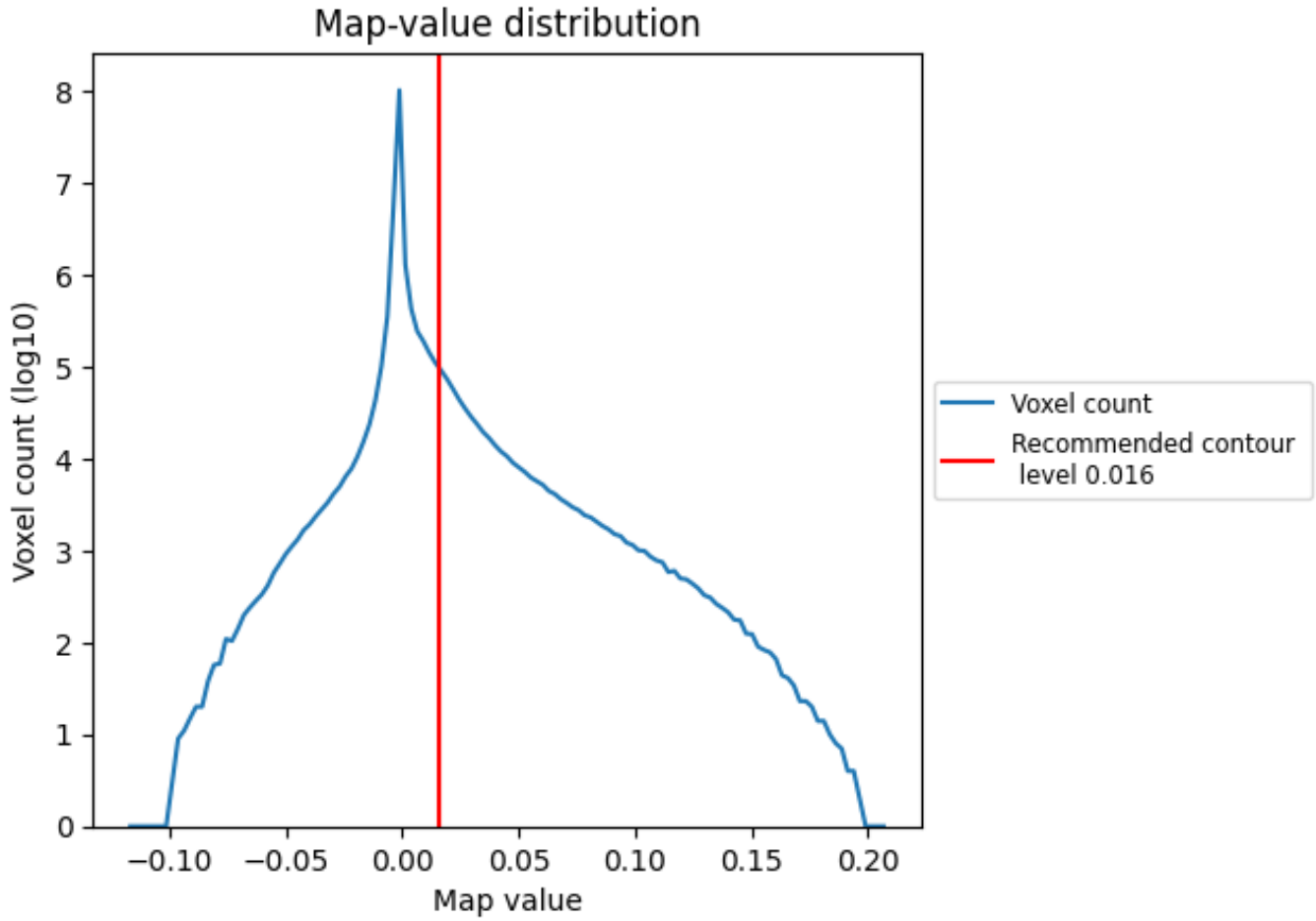
6.6 Mask visualisation (i)

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

7 Map analysis

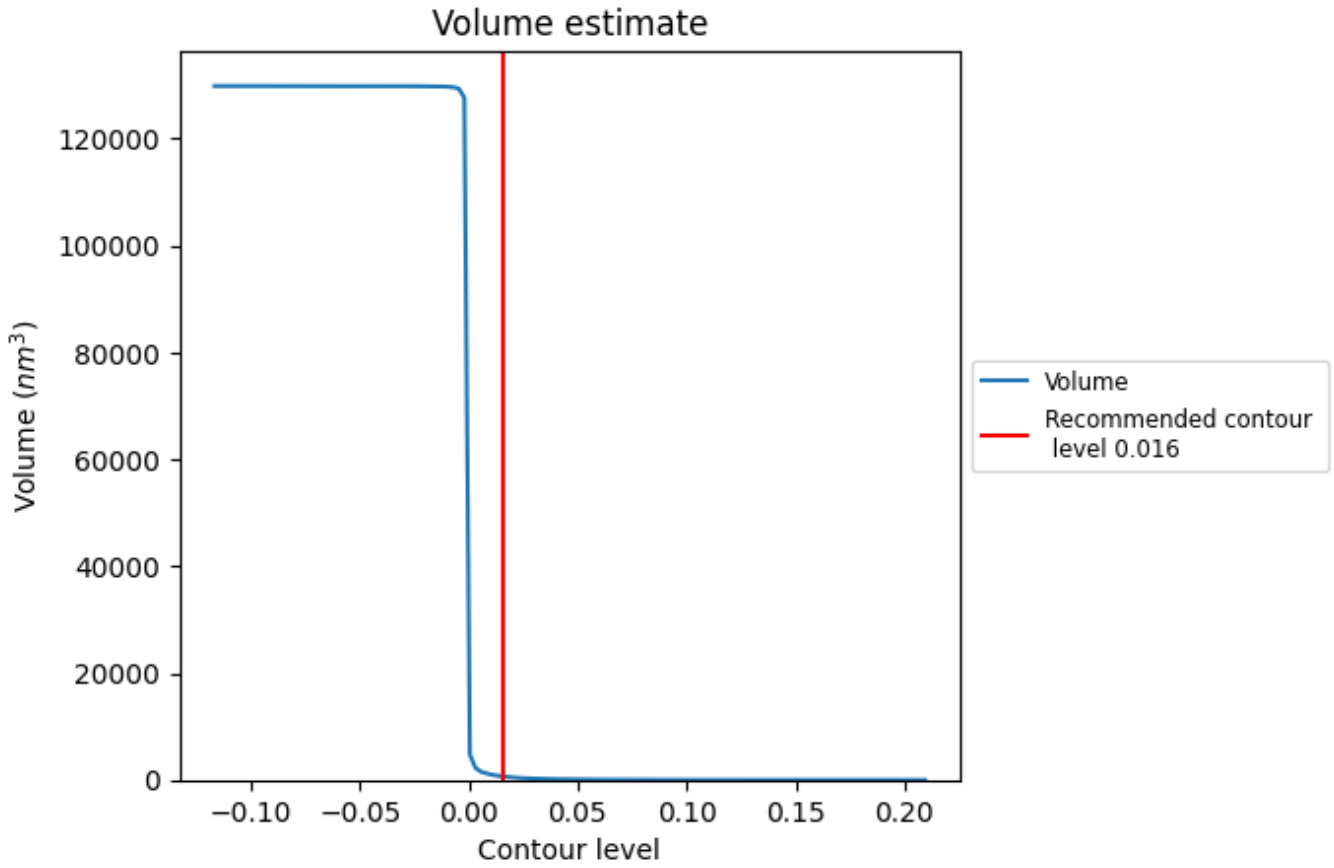
This section contains the results of statistical analysis of the map.

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.

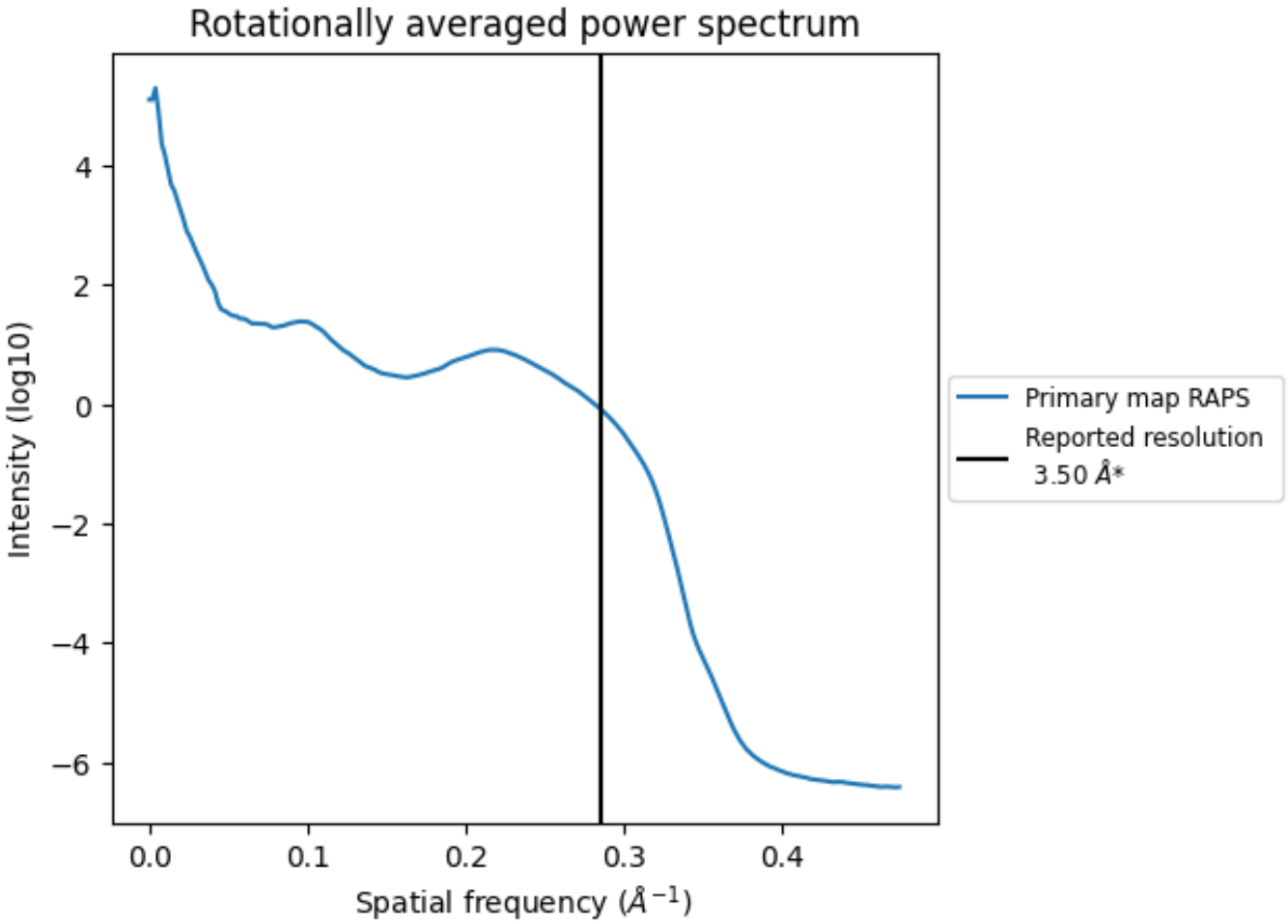
7.2 Volume estimate (i)



The volume at the recommended contour level is 642 nm^3; this corresponds to an approximate mass of 580 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 0.286 Å^{-1}

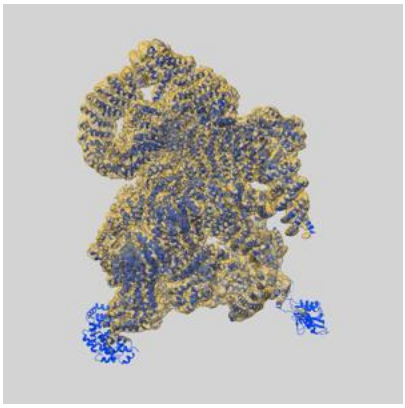
8 Fourier-Shell correlation

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

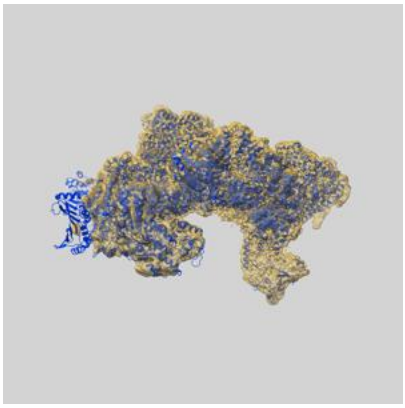
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-30473 and PDB model 7CUN. Per-residue inclusion information can be found in section 3 on page 7.

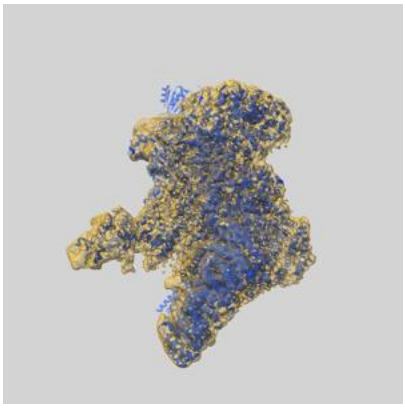
9.1 Map-model overlay



X



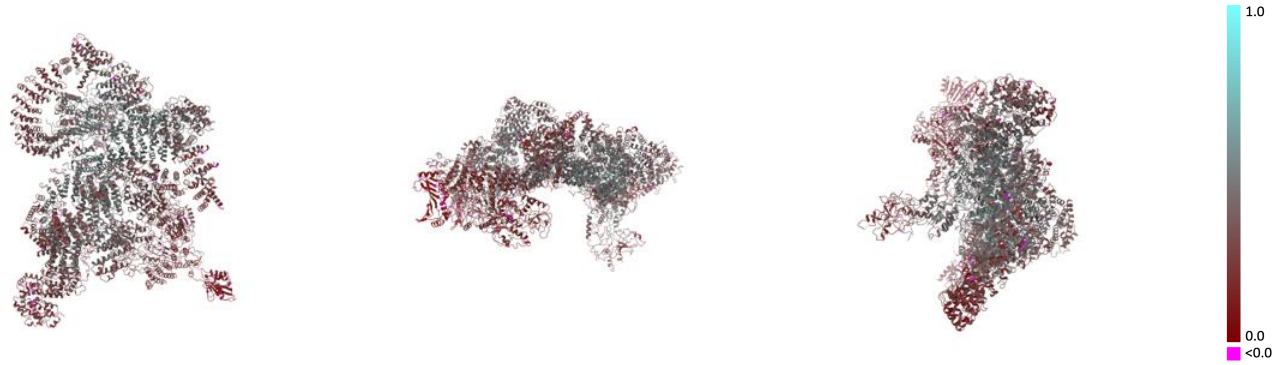
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.016 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.

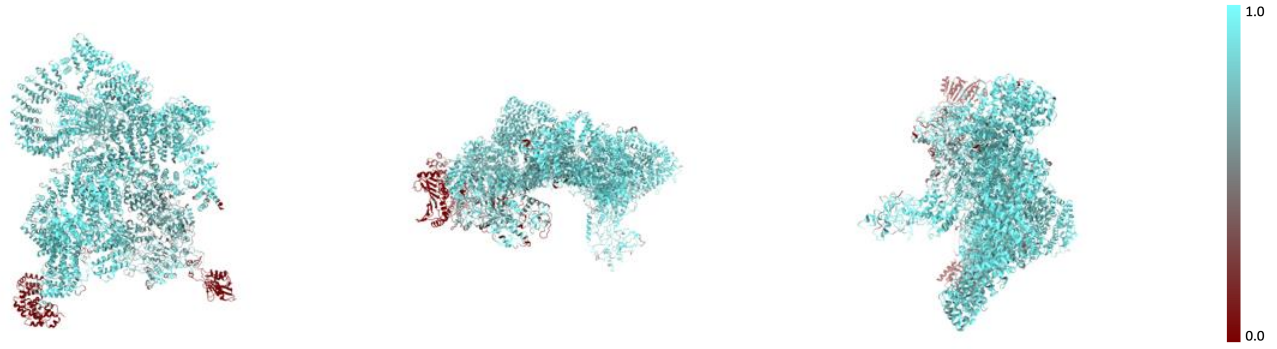
9.2 Q-score mapped to coordinate model (i)



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.

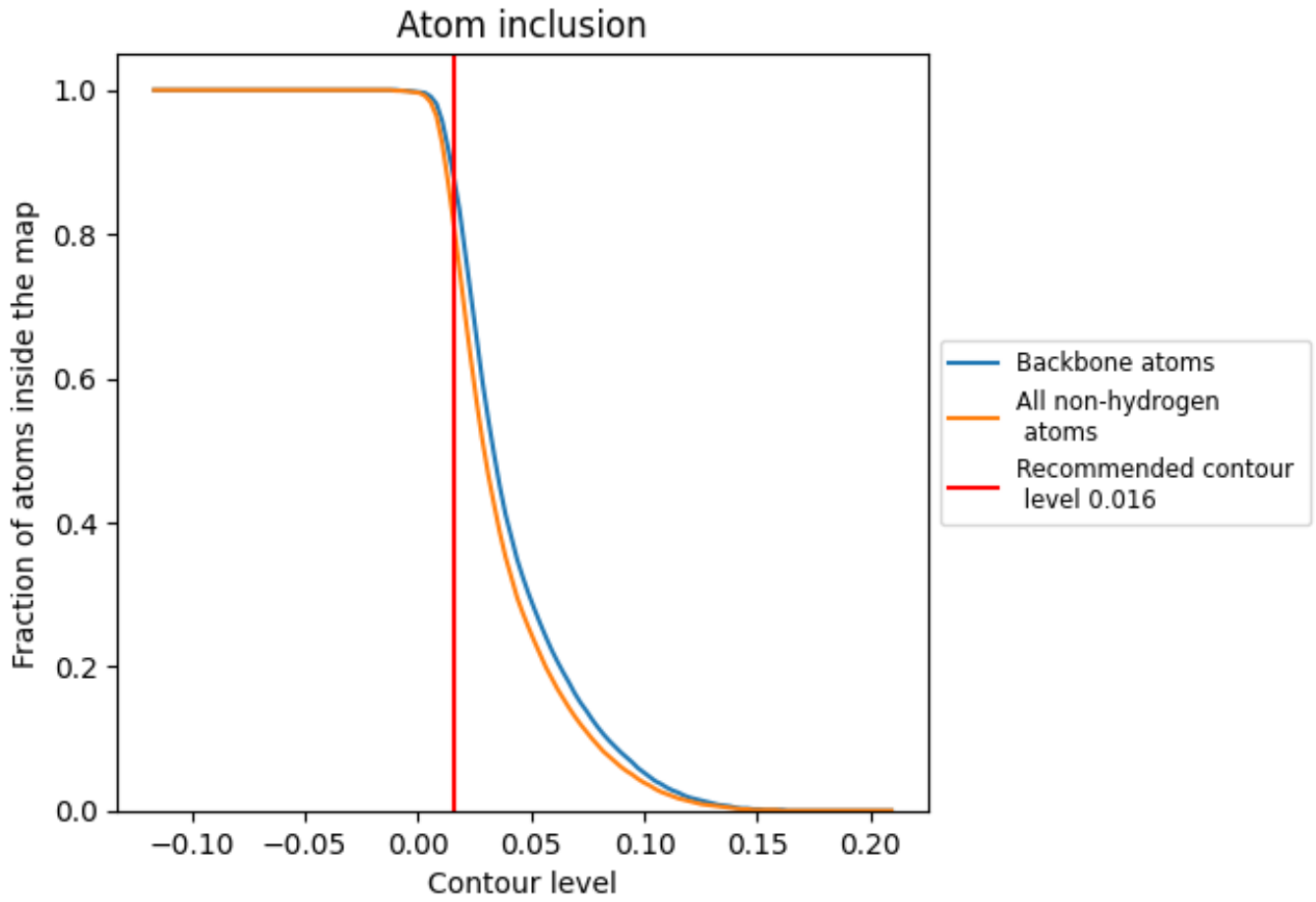
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.016).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 81% 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 (0.016) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8090	0.3500
A	0.8810	0.3750
B	0.8790	0.3980
D	0.8020	0.3170
E	0.9000	0.4070
F	0.8630	0.3560
G	0.8850	0.4230
H	0.8990	0.3900
I	0.6190	0.2110
K	0.5140	0.1700
P	0.8400	0.3000
Q	0.9310	0.5100
U	0.1300	0.2180