wwPDB EM Validation Summary Report

May 27, 2024 – 12:56 PM JST

PDB ID	:	7EGM
EMDB ID	:	EMD-31136
Title	:	The SRM module of SWI/SNF-nucleosome complex
Authors	:	Chen, Z.C.; Chen, K.J.; He, Z.Y.; Ye, Y.P.
Deposited on	:	2021-03-24
Resolution	:	3.60 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:

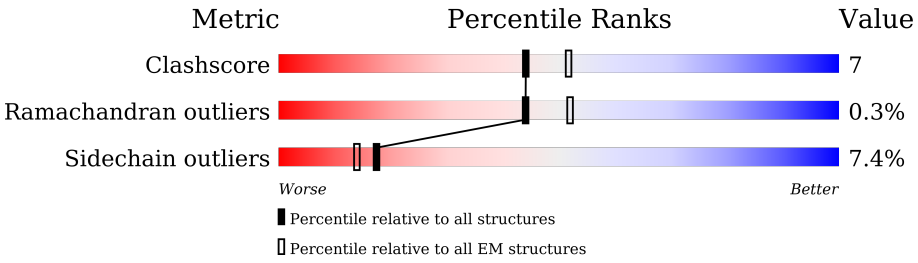
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.60 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	982	9% · 89%
2	B	1093	37% 8% · 54%
3	C	918	29% 7% · 63%
4	D	836	6% 40% 6% · 53%
4	E	836	8% 39% 11% · 50%
5	H	566	23% 54% 13% · 31%
6	I	332	25% 7% 68%
7	J	634	22% · 74%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 19819 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 Transcription regulatory protein SNF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	109	877	554	151	170	2	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1401	GLY	-	expression tag	UNP P22082
A	1402	GLY	-	expression tag	UNP P22082
A	1403	SER	-	expression tag	UNP P22082
A	1404	GLY	-	expression tag	UNP P22082
A	1405	GLY	-	expression tag	UNP P22082
A	1406	HIS	-	expression tag	UNP P22082
A	1407	HIS	-	expression tag	UNP P22082
A	1408	HIS	-	expression tag	UNP P22082
A	1409	HIS	-	expression tag	UNP P22082
A	1410	HIS	-	expression tag	UNP P22082
A	1411	HIS	-	expression tag	UNP P22082

- Molecule 2 is a protein called SWI/SNF chromatin-remodeling complex subunit SWI1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	507	4098	2663	672	752	11	0	0

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1315	GLY	-	expression tag	UNP P09547
B	1316	GLY	-	expression tag	UNP P09547
B	1317	SER	-	expression tag	UNP P09547
B	1318	GLY	-	expression tag	UNP P09547
B	1319	GLY	-	expression tag	UNP P09547
B	1320	TRP	-	expression tag	UNP P09547

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1321	SER	-	expression tag	UNP P09547
B	1322	HIS	-	expression tag	UNP P09547
B	1323	PRO	-	expression tag	UNP P09547
B	1324	GLN	-	expression tag	UNP P09547
B	1325	PHE	-	expression tag	UNP P09547
B	1326	GLU	-	expression tag	UNP P09547
B	1327	LYS	-	expression tag	UNP P09547
B	1328	TRP	-	expression tag	UNP P09547
B	1329	SER	-	expression tag	UNP P09547
B	1330	HIS	-	expression tag	UNP P09547
B	1331	PRO	-	expression tag	UNP P09547
B	1332	GLN	-	expression tag	UNP P09547
B	1333	PHE	-	expression tag	UNP P09547
B	1334	GLU	-	expression tag	UNP P09547
B	1335	LYS	-	expression tag	UNP P09547
B	1336	TRP	-	expression tag	UNP P09547
B	1337	SER	-	expression tag	UNP P09547
B	1338	HIS	-	expression tag	UNP P09547
B	1339	PRO	-	expression tag	UNP P09547
B	1340	GLN	-	expression tag	UNP P09547
B	1341	PHE	-	expression tag	UNP P09547
B	1342	GLU	-	expression tag	UNP P09547
B	1343	LYS	-	expression tag	UNP P09547

- Molecule 3 is a protein called SWI/SNF chromatin-remodeling complex subunit SNF5.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	337	Total	C	N	O	S	0	0
			2747	1724	471	542	10		

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	906	GLY	-	expression tag	UNP P18480
C	907	GLY	-	expression tag	UNP P18480
C	908	SER	-	expression tag	UNP P18480
C	909	GLY	-	expression tag	UNP P18480
C	910	GLY	-	expression tag	UNP P18480
C	911	ASP	-	expression tag	UNP P18480
C	912	TYR	-	expression tag	UNP P18480
C	913	LYS	-	expression tag	UNP P18480
C	914	ASP	-	expression tag	UNP P18480

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	915	ASP	-	expression tag	UNP P18480
C	916	ASP	-	expression tag	UNP P18480
C	917	ASP	-	expression tag	UNP P18480
C	918	LYS	-	expression tag	UNP P18480

- Molecule 4 is a protein called SWI/SNF complex subunit SWI3.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	395	Total	C	N	O	S	0	0
			3250	2082	560	597	11		
4	E	422	Total	C	N	O	S	0	0
			3454	2208	593	641	12		

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	826	GLY	-	expression tag	UNP P32591
D	827	GLY	-	expression tag	UNP P32591
D	828	SER	-	expression tag	UNP P32591
D	829	GLY	-	expression tag	UNP P32591
D	830	GLY	-	expression tag	UNP P32591
D	831	HIS	-	expression tag	UNP P32591
D	832	HIS	-	expression tag	UNP P32591
D	833	HIS	-	expression tag	UNP P32591
D	834	HIS	-	expression tag	UNP P32591
D	835	HIS	-	expression tag	UNP P32591
D	836	HIS	-	expression tag	UNP P32591
E	826	GLY	-	expression tag	UNP P32591
E	827	GLY	-	expression tag	UNP P32591
E	828	SER	-	expression tag	UNP P32591
E	829	GLY	-	expression tag	UNP P32591
E	830	GLY	-	expression tag	UNP P32591
E	831	HIS	-	expression tag	UNP P32591
E	832	HIS	-	expression tag	UNP P32591
E	833	HIS	-	expression tag	UNP P32591
E	834	HIS	-	expression tag	UNP P32591
E	835	HIS	-	expression tag	UNP P32591
E	836	HIS	-	expression tag	UNP P32591

- Molecule 5 is a protein called Transcription regulatory protein SNF12.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	H	388	Total	C	N	O	S	0	0
			3145	1995	536	603	11		

- Molecule 6 is a protein called Transcription regulatory protein SNF6.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	I	106	Total	C	N	O	S	0	0
			888	553	158	173	4		

- Molecule 7 is a protein called SWI/SNF global transcription activator complex subunit SWP82.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	J	167	Total	C	N	O	S	0	0
			1360	877	241	240	2		

There are 11 discrepancies between the modelled and reference sequences:

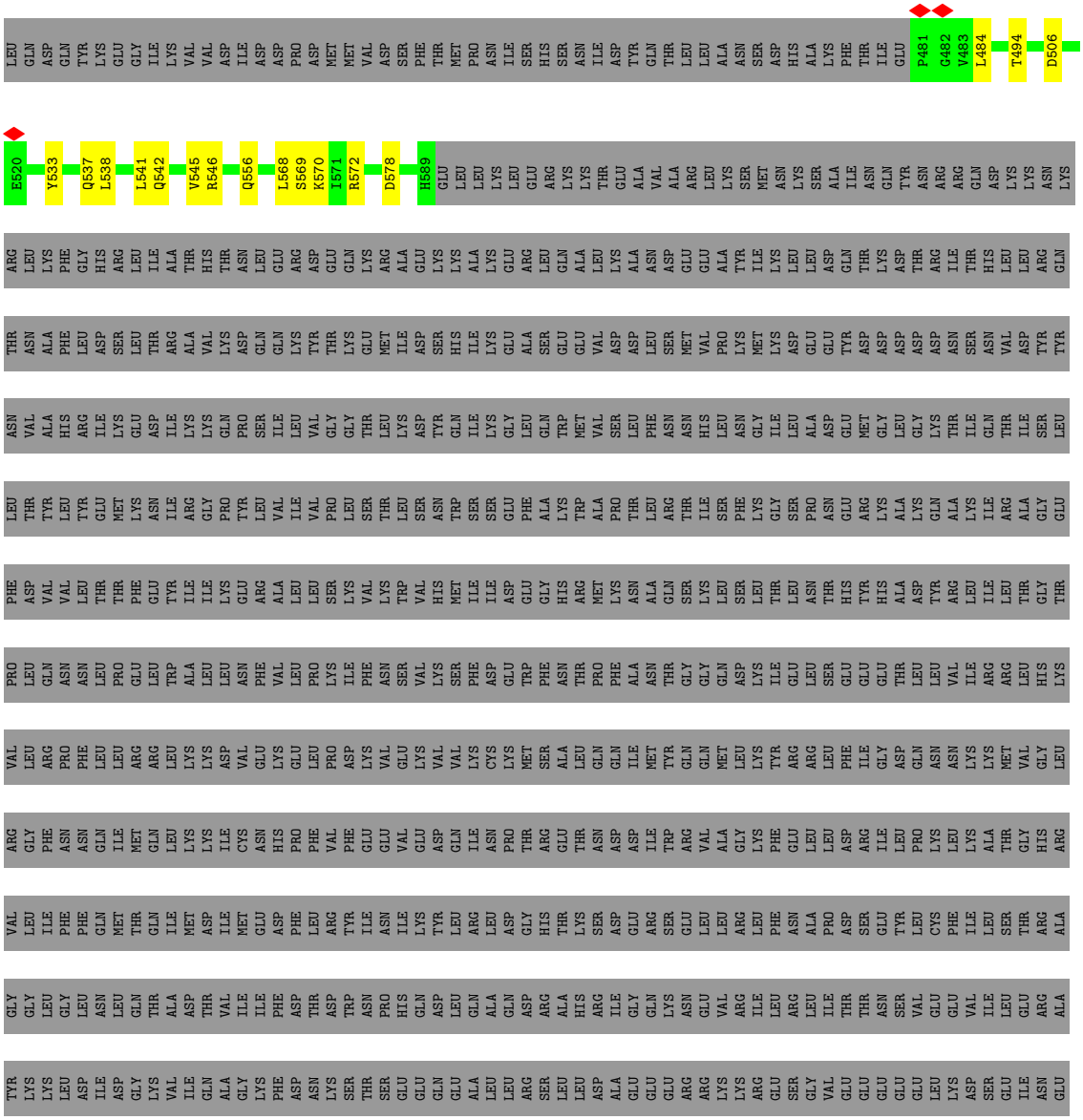
Chain	Residue	Modelled	Actual	Comment	Reference
J	624	GLY	-	expression tag	UNP P43554
J	625	GLY	-	expression tag	UNP P43554
J	626	SER	-	expression tag	UNP P43554
J	627	GLY	-	expression tag	UNP P43554
J	628	GLY	-	expression tag	UNP P43554
J	629	HIS	-	expression tag	UNP P43554
J	630	HIS	-	expression tag	UNP P43554
J	631	HIS	-	expression tag	UNP P43554
J	632	HIS	-	expression tag	UNP P43554
J	633	HIS	-	expression tag	UNP P43554
J	634	HIS	-	expression tag	UNP P43554

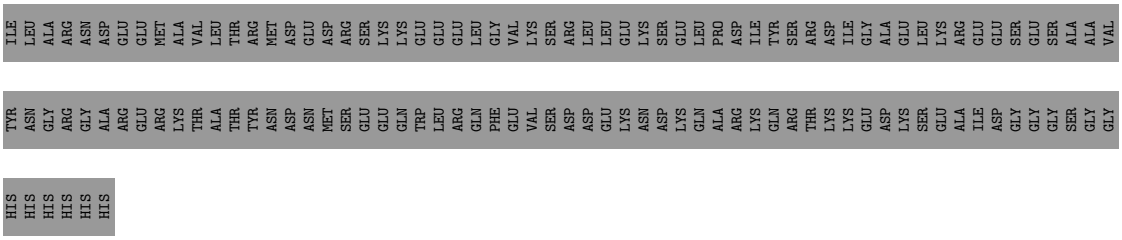
3 Residue-property plots (i)

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: Transcription regulatory protein SNF2

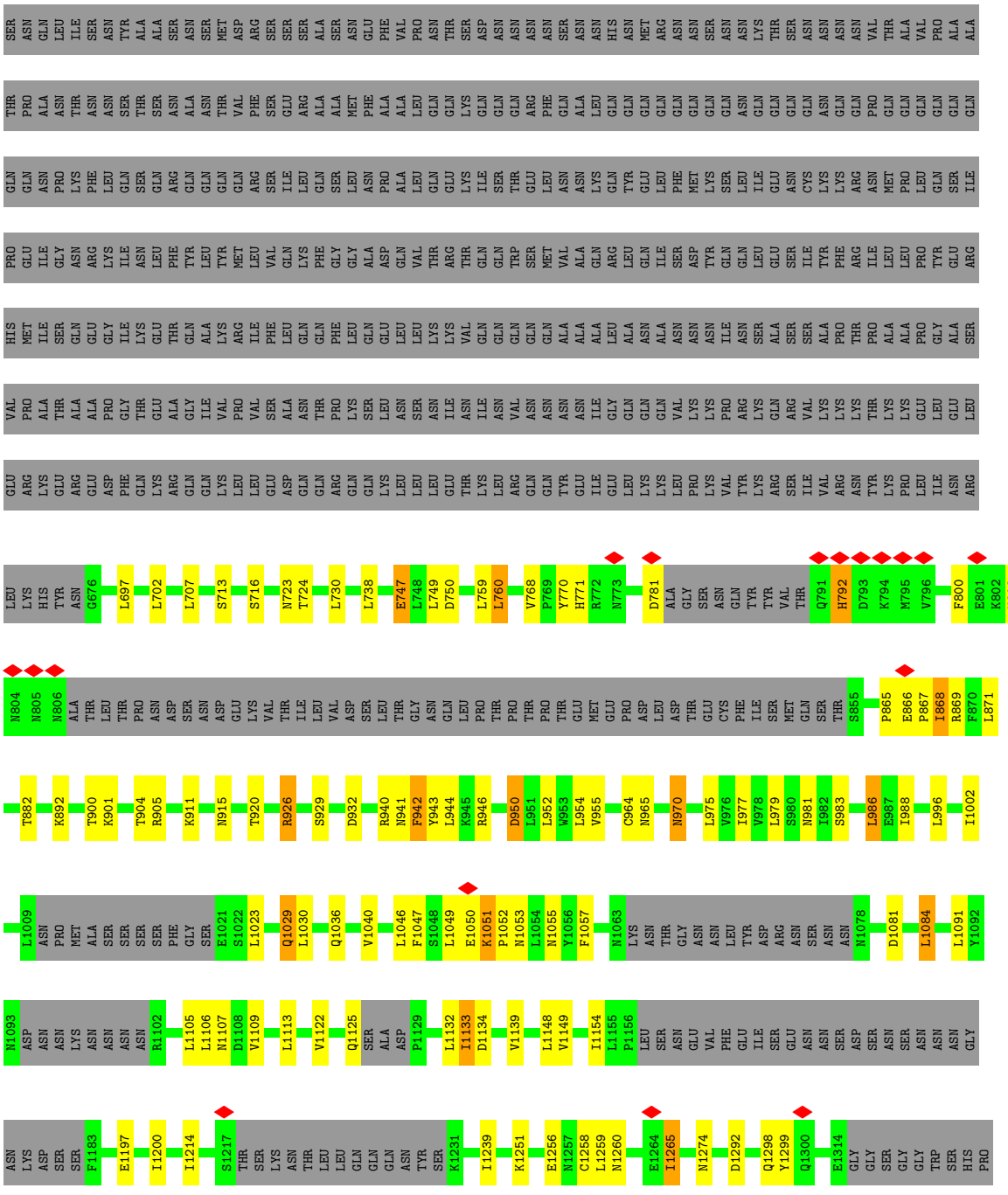
Chain A: 9% 89%





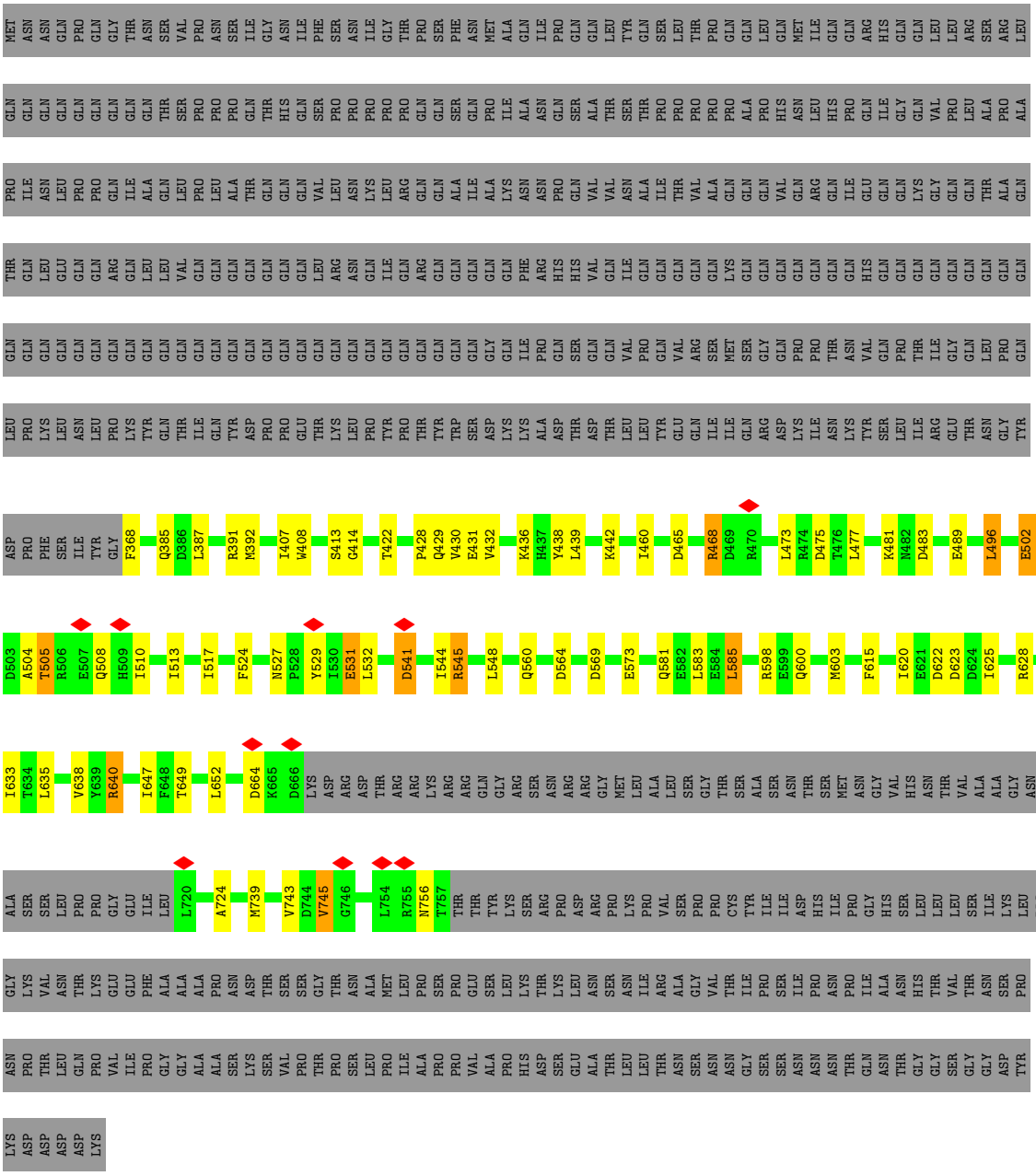
• Molecule 2: SWI/SNF chromatin-remodeling complex subunit SWI1

Chain B:



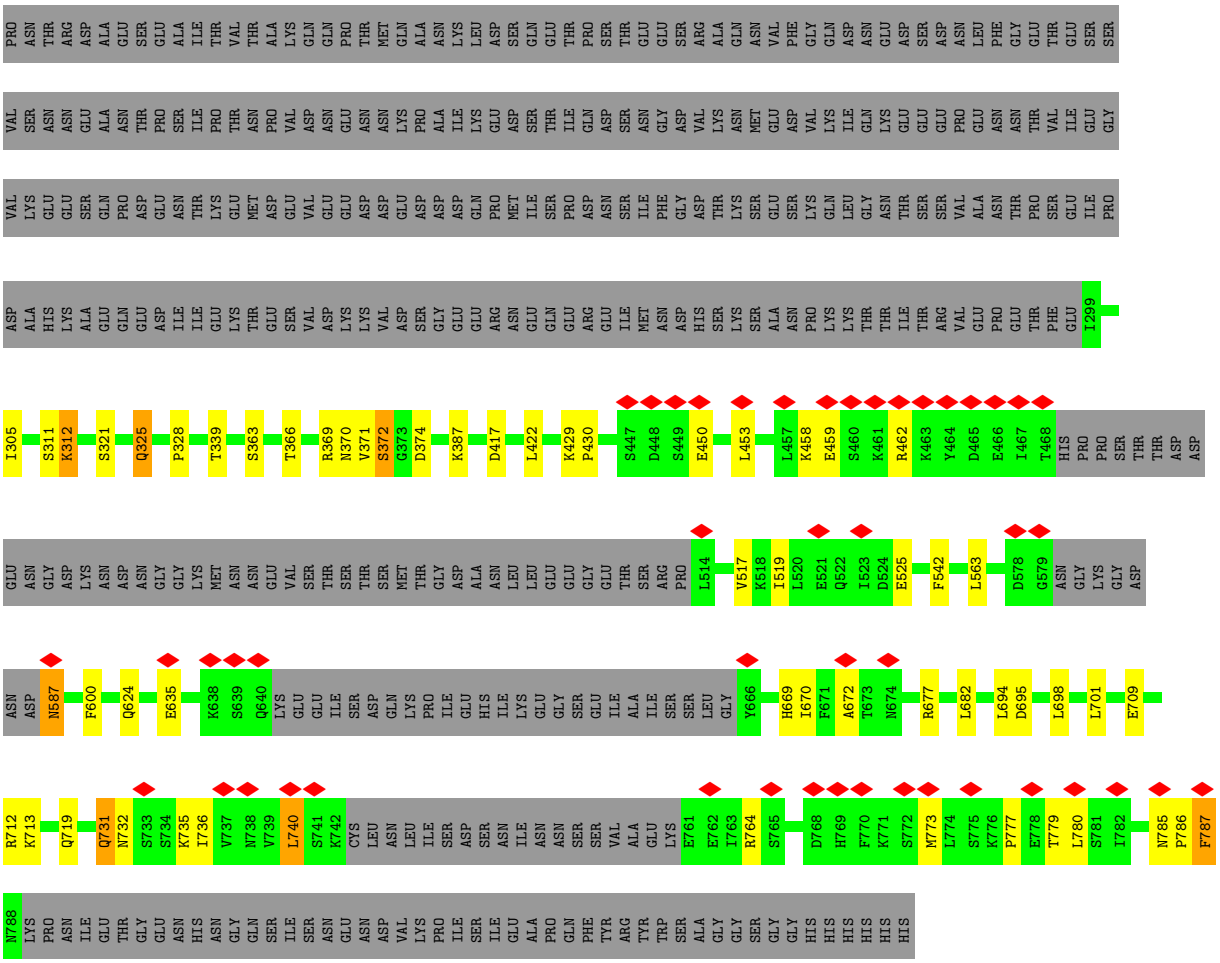
• Molecule 3: SWI/SNF chromatin-remodeling complex subunit SNF5

Chain C:



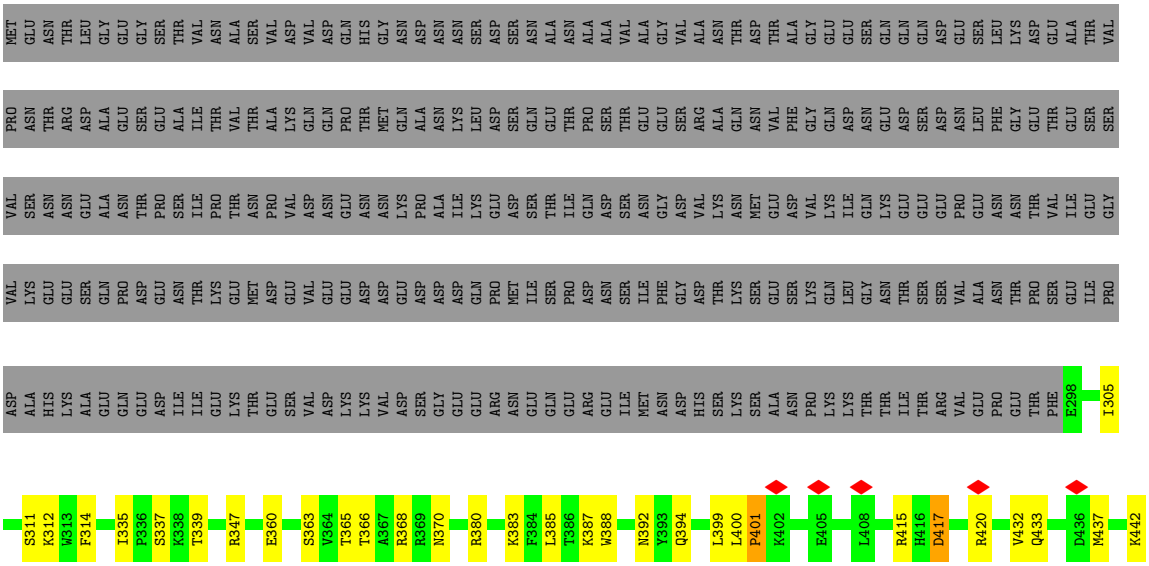
• Molecule 4: SWI/SNF complex subunit SWI3

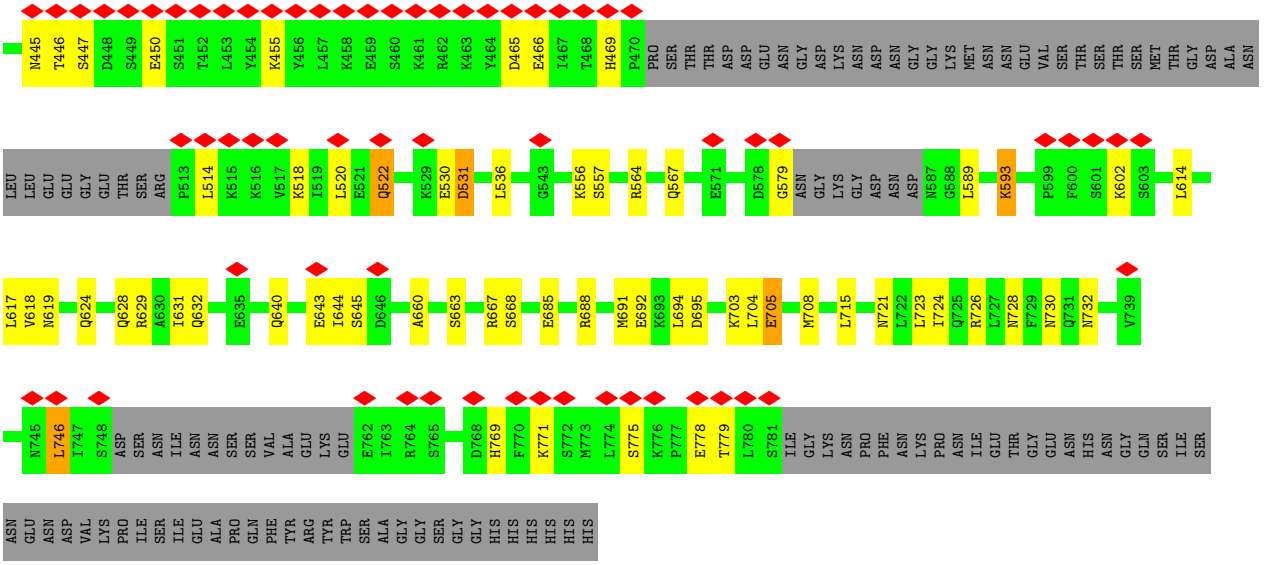
Chain D:



• Molecule 4: SWI/SNF complex subunit SWI3

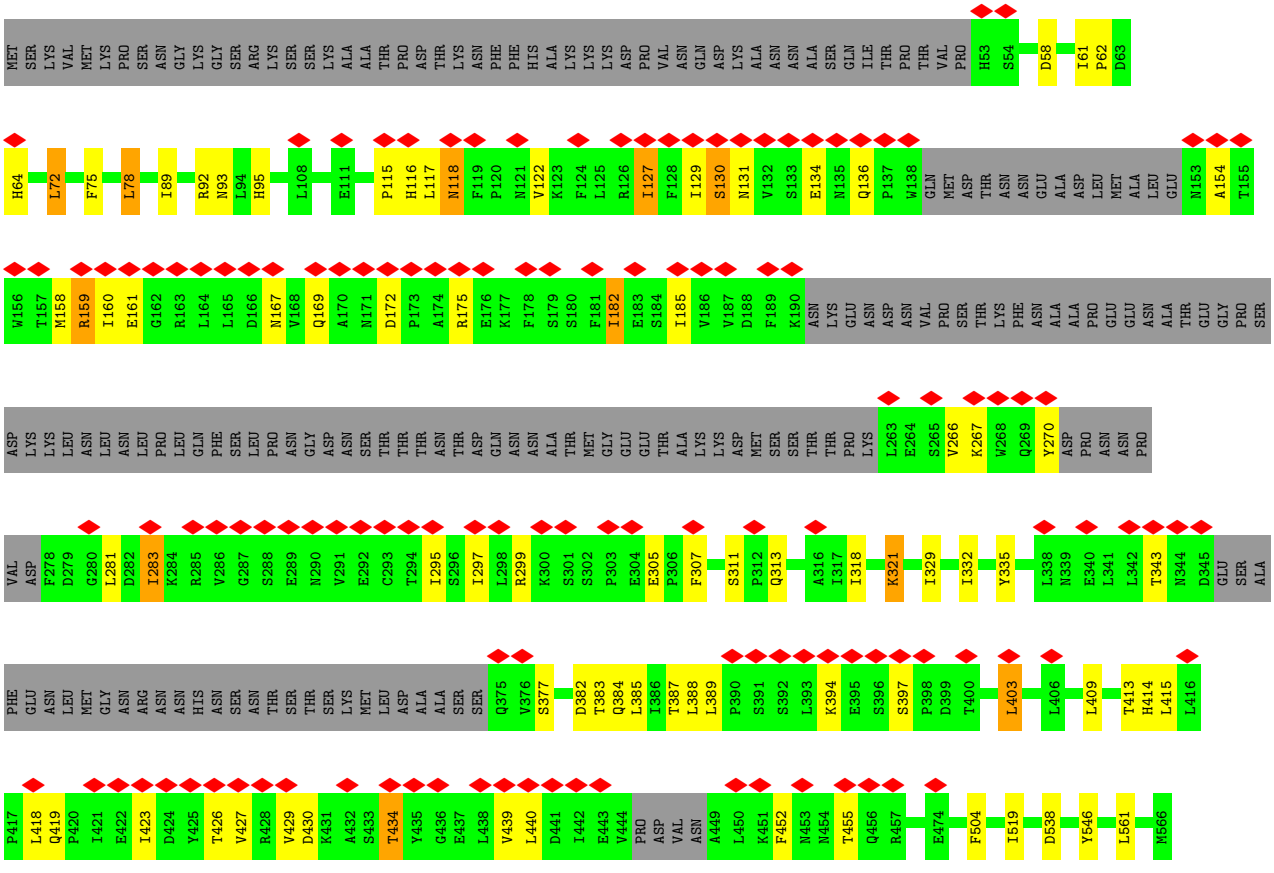
Chain E:





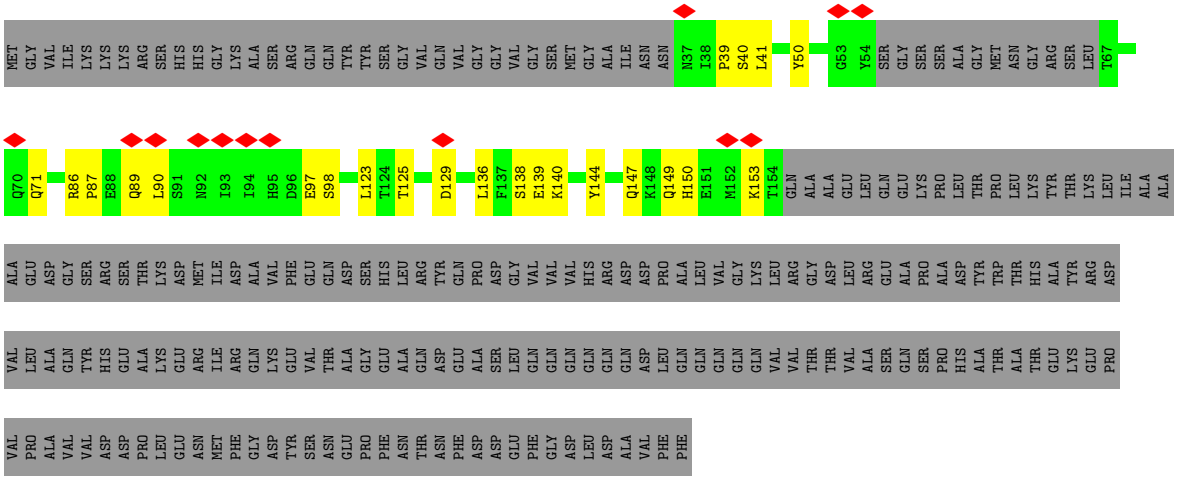
• Molecule 5: Transcription regulatory protein SNF12

Chain H:



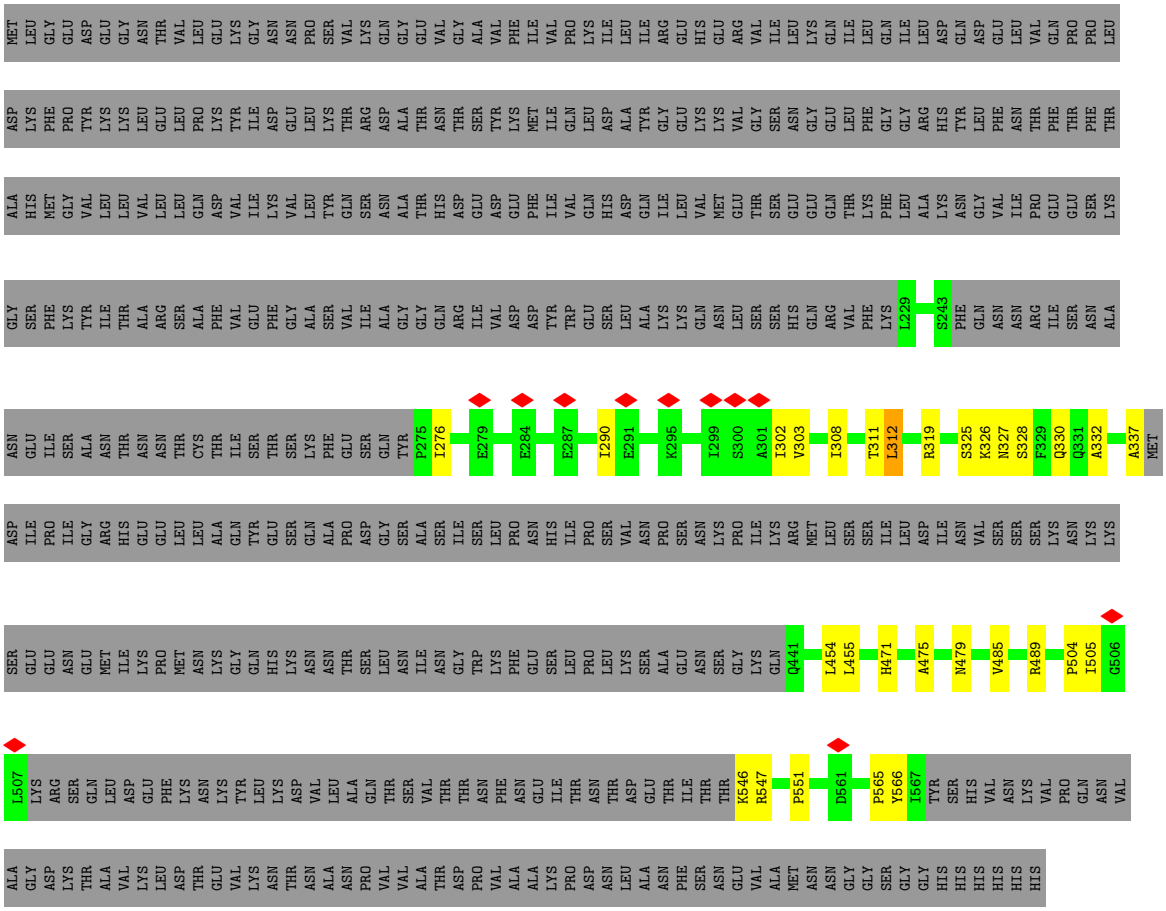
• Molecule 6: Transcription regulatory protein SNF6

Chain I:



• Molecule 7: SWI/SNF global transcription activator complex subunit SWP82

Chain J:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	210422	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	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.063	Depositor
Minimum map value	-0.028	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.0125	Depositor
Map size (Å)	429.68002, 429.68002, 429.68002	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.0742, 1.0742, 1.0742	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	0.28	0/894	0.44	0/1222
2	B	0.30	0/4171	0.49	0/5648
3	C	0.30	0/2800	0.49	0/3790
4	D	0.29	0/3318	0.45	0/4465
4	E	0.29	0/3526	0.47	0/4747
5	H	0.27	0/3201	0.46	0/4322
6	I	0.25	0/902	0.48	0/1211
7	J	0.27	0/1390	0.47	0/1874
All	All	0.29	0/20202	0.47	0/27279

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	877	0	864	21	0
2	B	4098	0	4239	93	0
3	C	2747	0	2670	39	0
4	D	3250	0	3302	33	0
4	E	3454	0	3510	49	0
5	H	3145	0	3143	57	0
6	I	888	0	867	9	0
7	J	1360	0	1400	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	19819	0	19995	275	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 275 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)
2:B:866:GLU:HG3	2:B:871:LEU:CD2	1.68	1.21
2:B:866:GLU:CG	2:B:871:LEU:HD21	1.71	1.19
1:A:546:ARG:HD3	5:H:546:TYR:CZ	1.79	1.17
2:B:1051:LYS:CB	2:B:1052:PRO:HD3	1.76	1.16
2:B:1051:LYS:HB3	2:B:1052:PRO:HD3	1.32	1.08

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	107/982 (11%)	102 (95%)	5 (5%)	0	100	100
2	B	489/1093 (45%)	461 (94%)	25 (5%)	3 (1%)	25	64
3	C	333/918 (36%)	299 (90%)	34 (10%)	0	100	100
4	D	385/836 (46%)	372 (97%)	13 (3%)	0	100	100
4	E	414/836 (50%)	389 (94%)	24 (6%)	1 (0%)	47	79
5	H	376/566 (66%)	346 (92%)	29 (8%)	1 (0%)	41	75
6	I	102/332 (31%)	93 (91%)	8 (8%)	1 (1%)	15	55
7	J	159/634 (25%)	147 (92%)	11 (7%)	1 (1%)	25	64
All	All	2365/6197 (38%)	2209 (93%)	149 (6%)	7 (0%)	44	75

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	868	ILE
2	B	942	PHE
2	B	1051	LYS
5	H	62	PRO
7	J	504	PRO

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	100/889 (11%)	96 (96%)	4 (4%)	31	65
2	B	484/1017 (48%)	448 (93%)	36 (7%)	13	46
3	C	308/832 (37%)	281 (91%)	27 (9%)	10	40
4	D	368/758 (48%)	342 (93%)	26 (7%)	14	48
4	E	393/758 (52%)	363 (92%)	30 (8%)	13	45
5	H	360/517 (70%)	330 (92%)	30 (8%)	11	42
6	I	102/288 (35%)	93 (91%)	9 (9%)	10	40
7	J	150/565 (26%)	144 (96%)	6 (4%)	31	65
All	All	2265/5624 (40%)	2097 (93%)	168 (7%)	17	46

5 of 168 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	H	418	LEU
4	E	415	ARG
5	H	434	THR
6	I	140	LYS
4	E	522	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	785	ASN
6	I	48	ASN
4	E	730	ASN
5	H	93	ASN
5	H	453	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.

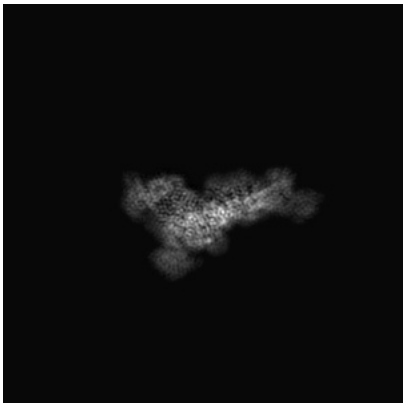
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-31136. These allow visual inspection of the internal detail of the map and identification of artifacts.

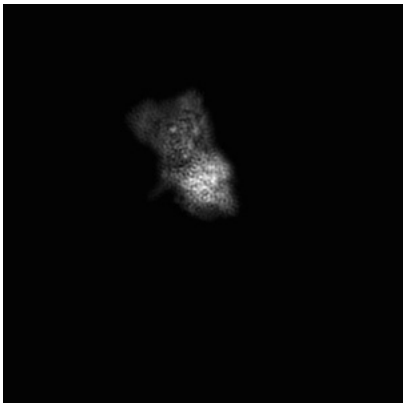
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

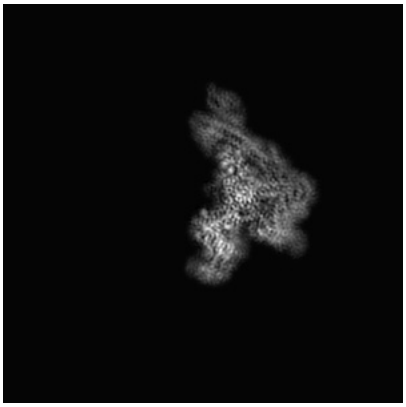
6.1.1 Primary map



X



Y

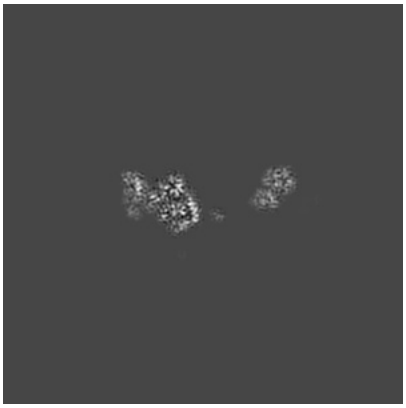


Z

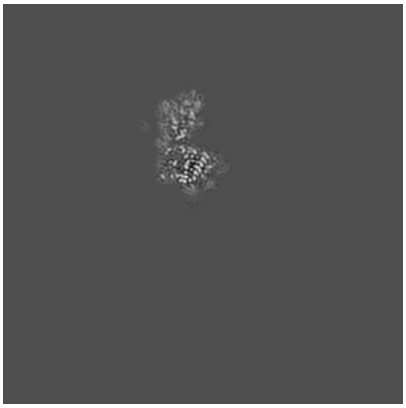
The images above show the map projected in three orthogonal directions.

6.2 Central slices

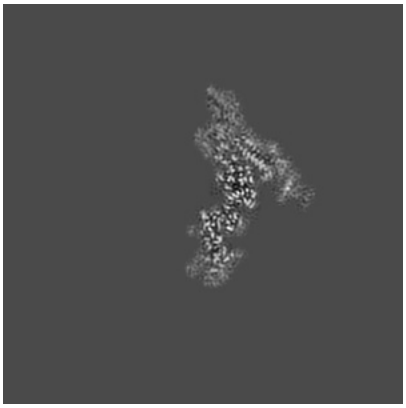
6.2.1 Primary map



X Index: 200



Y Index: 200

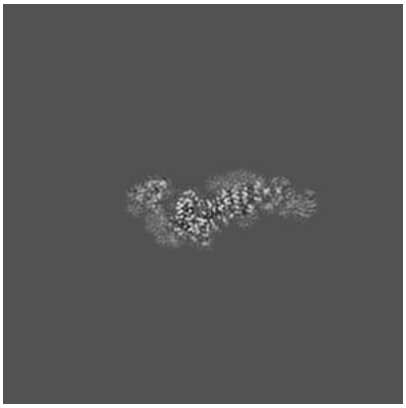


Z Index: 200

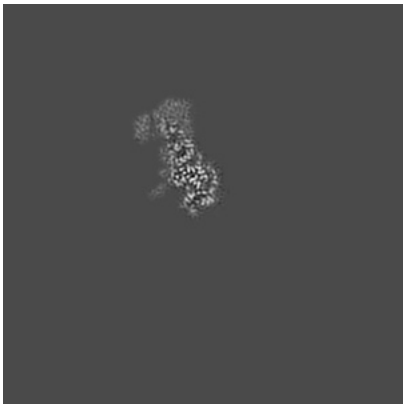
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

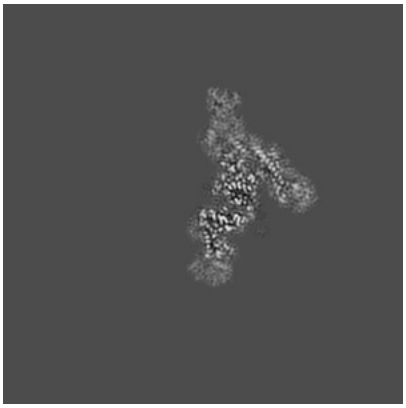
6.3.1 Primary map



X Index: 226



Y Index: 186

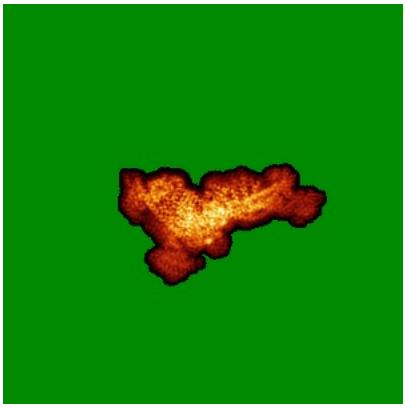


Z Index: 195

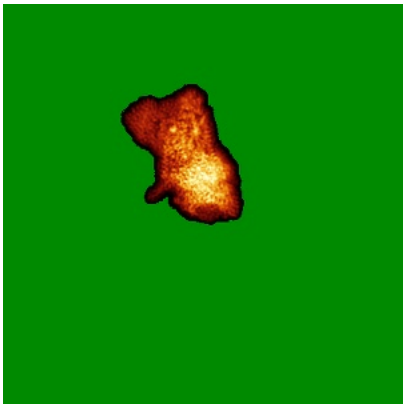
The images above show the largest variance slices of the map in three orthogonal directions.

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

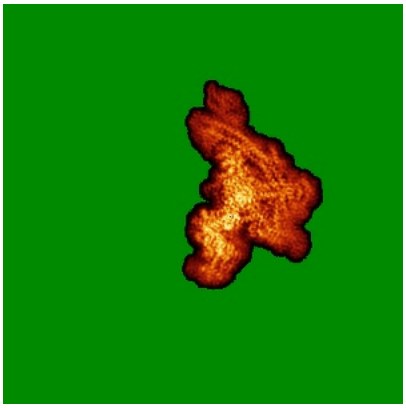
6.4.1 Primary map



X



Y

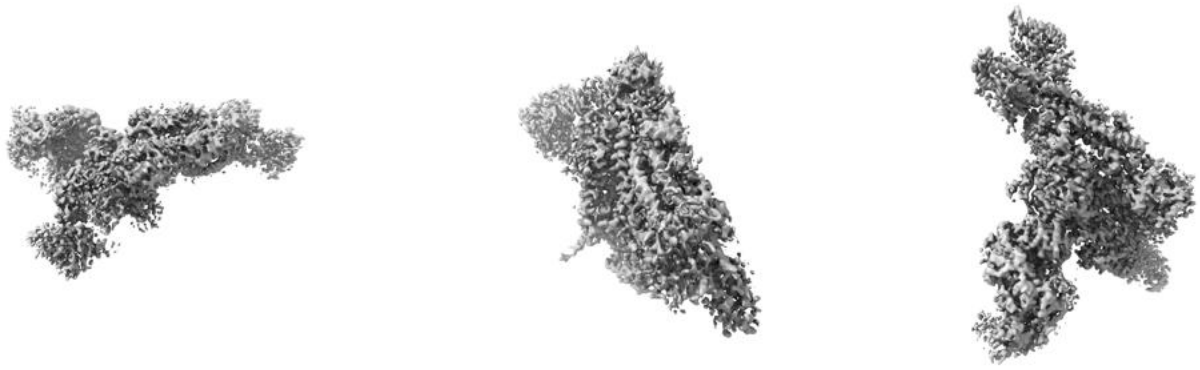


Z

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 0.0125. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

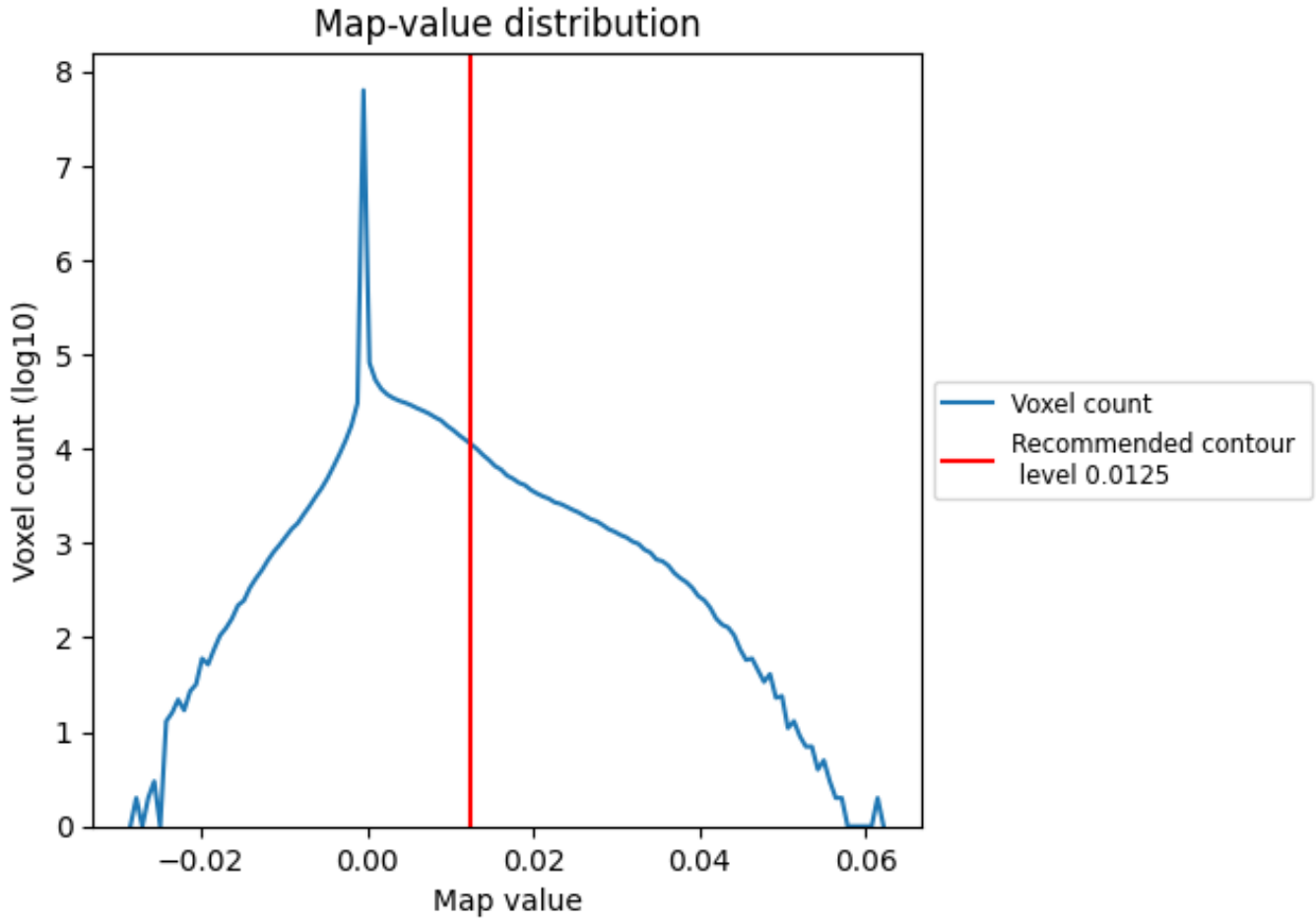
6.6 Mask visualisation (i)

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

7 Map analysis

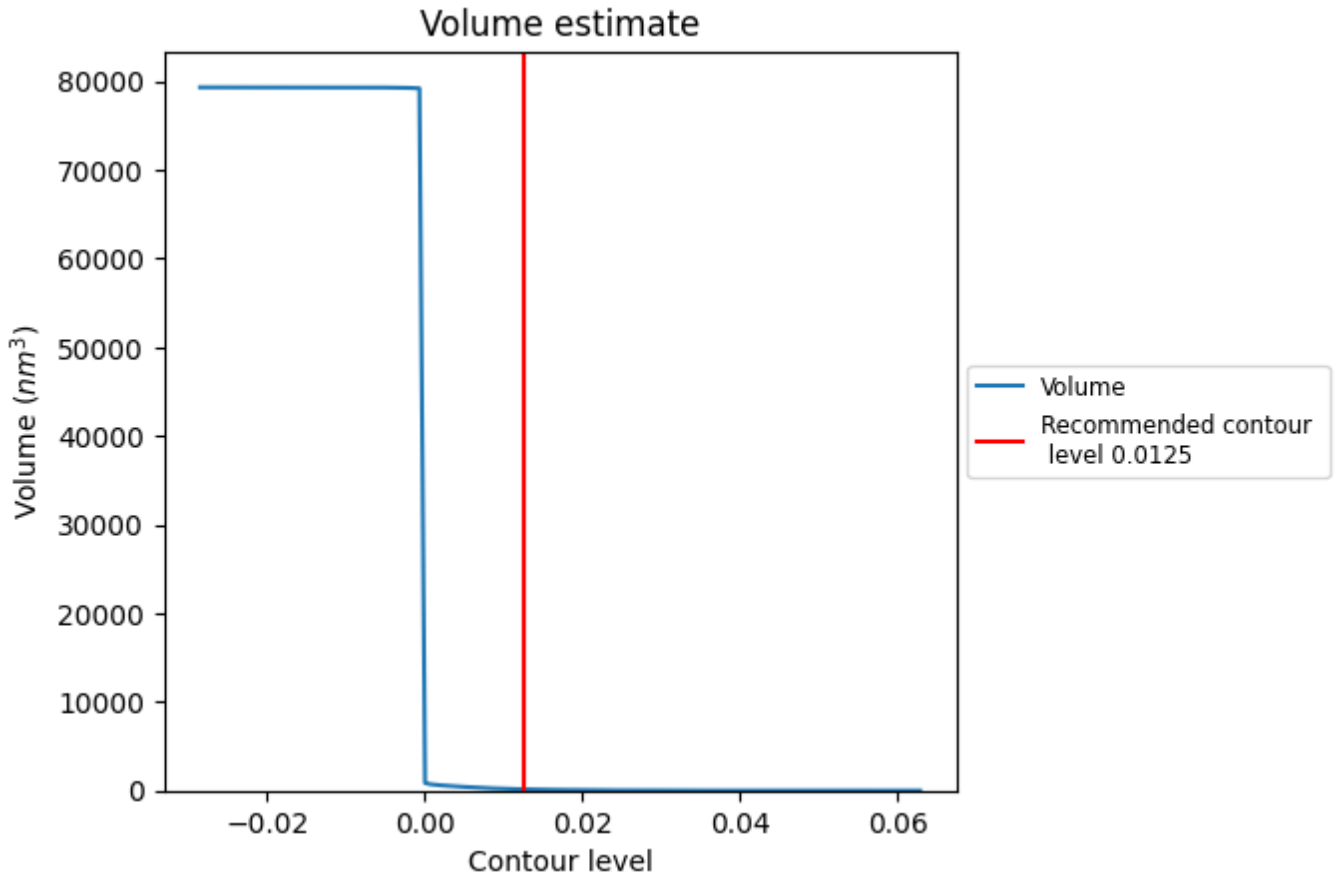
This section contains the results of statistical analysis of the map.

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.

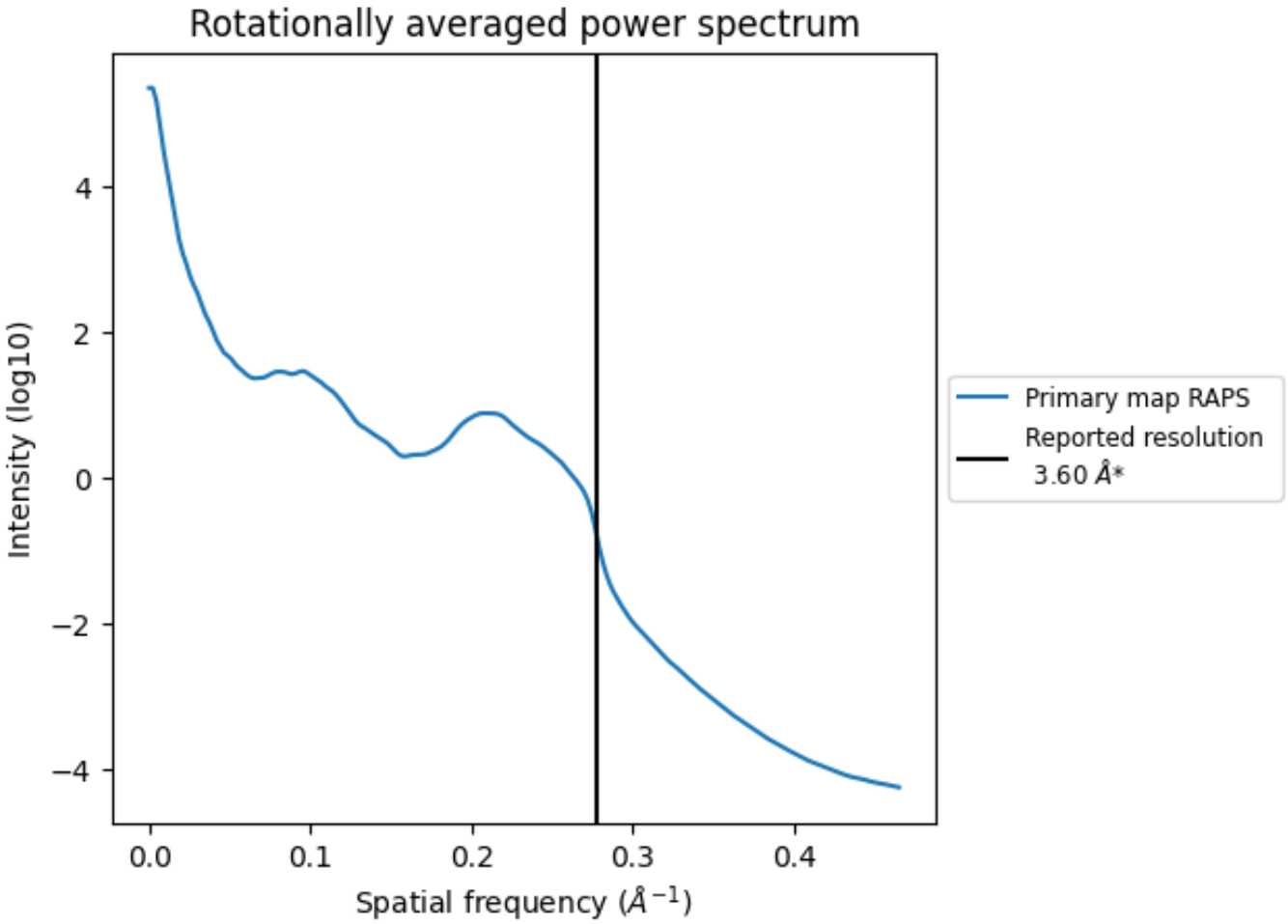
7.2 Volume estimate (i)



The volume at the recommended contour level is 143 nm^3; this corresponds to an approximate mass of 129 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 0.278 $Å^{-1}$

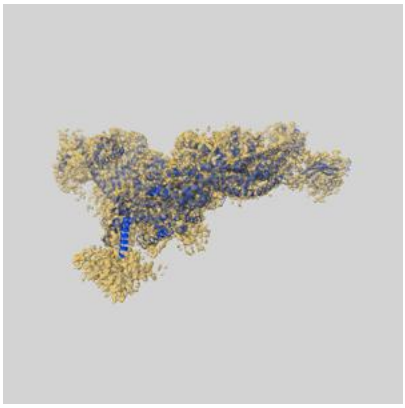
8 Fourier-Shell correlation

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

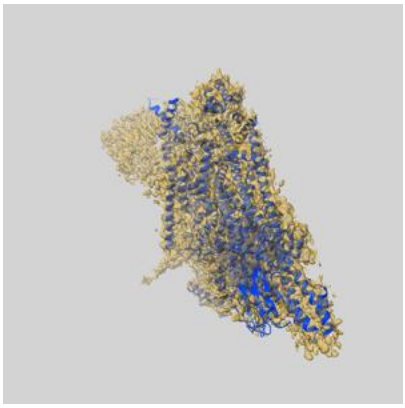
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-31136 and PDB model 7EGM. Per-residue inclusion information can be found in section 3 on page 7.

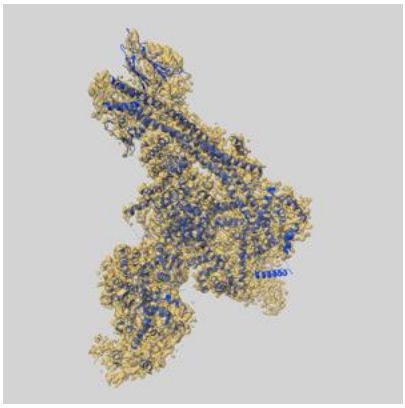
9.1 Map-model overlay



X



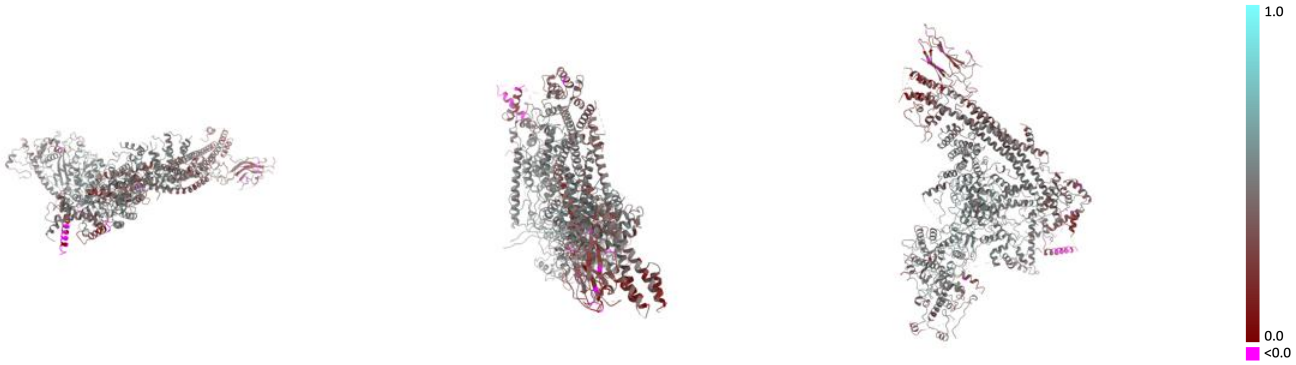
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0125 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.

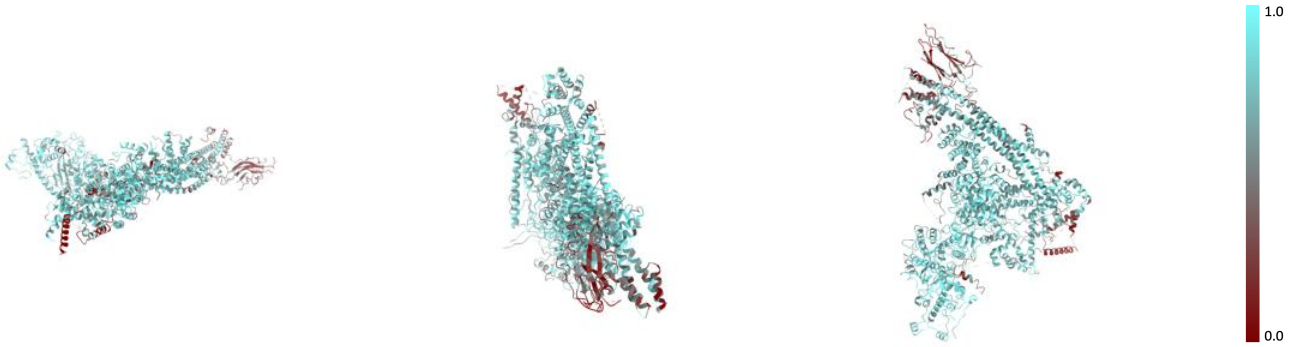
9.2 Q-score mapped to coordinate model (i)



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.

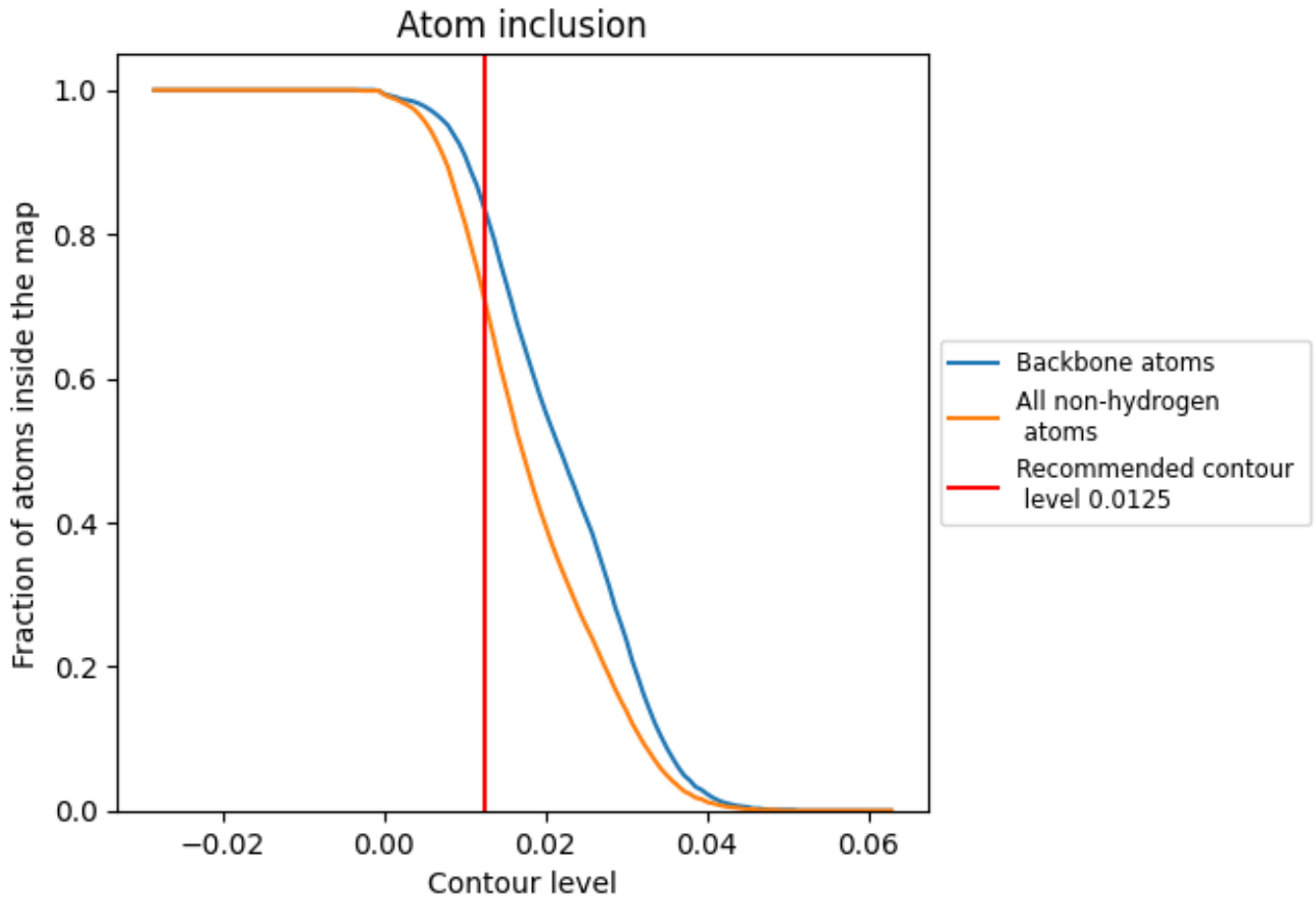
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0125).

9.4 Atom inclusion (i)



At the recommended contour level, 83% of all backbone atoms, 71% 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 (0.0125) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7060	0.4320
A	0.7900	0.4760
B	0.7950	0.4830
C	0.7860	0.4710
D	0.7110	0.4180
E	0.6730	0.3910
H	0.5340	0.3670
I	0.6310	0.4400
J	0.7400	0.4530