wwPDB EM Validation Summary Report

Nov 3, 2024 – 03:55 AM EST

PDB ID	:	6XF8
EMDB ID	:	EMD-22166
Title	:	DLP 5 fold
Authors	:	Sutton, G.; Sun, D.P.; Fu, X.F.; Kotecha, A.; Hecksel, G.W.; Clare, D.K.; Zhang, P.; Stuart, D.; Boyce, M.
Deposited on	:	2020-06-15
Resolution	:	6.50 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:

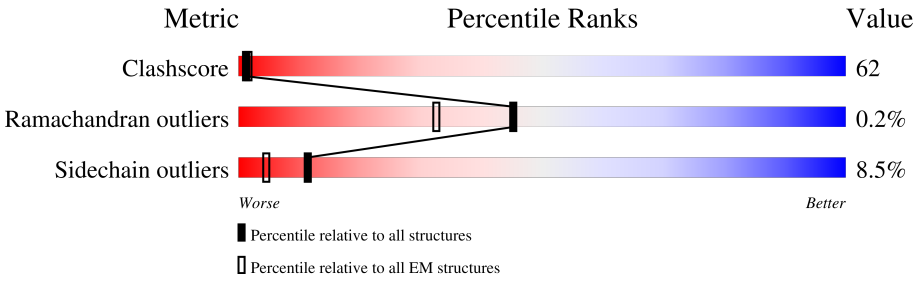
EMDB validation analysis	:	0.0.1.dev113
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	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 6.50 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.



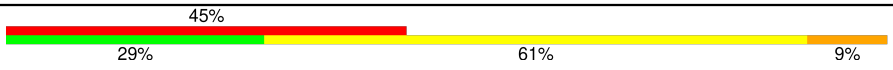
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	F	633	58% 40% 54%
1	K	633	61% 27% 61% 8%
2	G	365	42% 34% 59% 7%
2	H	365	52% 42% 56%
2	I	365	36% 45% 53%
3	E	417	53% 36% 58% 7%
4	B	1059	55% 37% 56% 5%
4	C	1059	50% 35% 58% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	A	1288	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 47830 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 Outer capsid protein mu-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	K	608	Total 4641	C 2950	N 770	O 903	S 18	0	0
1	F	608	Total 4641	C 2950	N 770	O 903	S 18	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	344	LEU	PRO	conflict	UNP P11077
K	359	PHE	LEU	conflict	UNP P11077
F	344	LEU	PRO	conflict	UNP P11077
F	359	PHE	LEU	conflict	UNP P11077

- Molecule 2 is a protein called Outer capsid protein sigma-3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	I	365	Total 2885	C 1818	N 508	O 531	S 28	0	0
2	H	365	Total 2885	C 1818	N 508	O 531	S 28	0	0
2	G	365	Total 2885	C 1818	N 508	O 531	S 28	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	104	CYS	ALA	conflict	UNP P07939
I	325	ASN	ASP	conflict	UNP P07939
H	104	CYS	ALA	conflict	UNP P07939
H	325	ASN	ASP	conflict	UNP P07939
G	104	CYS	ALA	conflict	UNP P07939
G	325	ASN	ASP	conflict	UNP P07939

- Molecule 3 is a protein called Inner capsid protein sigma-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	E	417	Total	C	N	O	S	0	0
			3313	2092	600	604	17		

- Molecule 4 is a protein called Inner capsid protein lambda-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	C	1051	Total	C	N	O	S	0	0
			8311	5314	1407	1540	50		
4	B	1031	Total	C	N	O	S	0	0
			8143	5208	1375	1510	50		

- Molecule 5 is a protein called mRNA (guanine-N(7)-)-methyltransferase.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	A	1284	Total	C	N	O	S	0	0
			10126	6468	1699	1917	42		

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	ASN	THR	conflict	UNP Q91RA5
A	78	ALA	VAL	conflict	UNP Q91RA5
A	97	VAL	ILE	conflict	UNP Q91RA5
A	108	ASN	SER	conflict	UNP Q91RA5
A	151	ALA	VAL	conflict	UNP Q91RA5
A	231	HIS	ASN	conflict	UNP Q91RA5
A	277	THR	ALA	conflict	UNP Q91RA5
A	365	THR	ALA	conflict	UNP Q91RA5
A	370	VAL	ILE	conflict	UNP Q91RA5
A	407	MET	VAL	conflict	UNP Q91RA5
A	425	ASP	GLU	conflict	UNP Q91RA5
A	508	SER	ALA	conflict	UNP Q91RA5
A	509	ARG	GLY	conflict	UNP Q91RA5
A	559	ILE	VAL	conflict	UNP Q91RA5
A	599	LEU	PHE	conflict	UNP Q91RA5
A	622	PRO	THR	conflict	UNP Q91RA5
A	661	HIS	GLN	conflict	UNP Q91RA5
A	775	ILE	VAL	conflict	UNP Q91RA5
A	799	SER	ALA	conflict	UNP Q91RA5
A	889	MET	ILE	conflict	UNP Q91RA5
A	916	LYS	ARG	conflict	UNP Q91RA5

Continued on next page...

Continued from previous page...

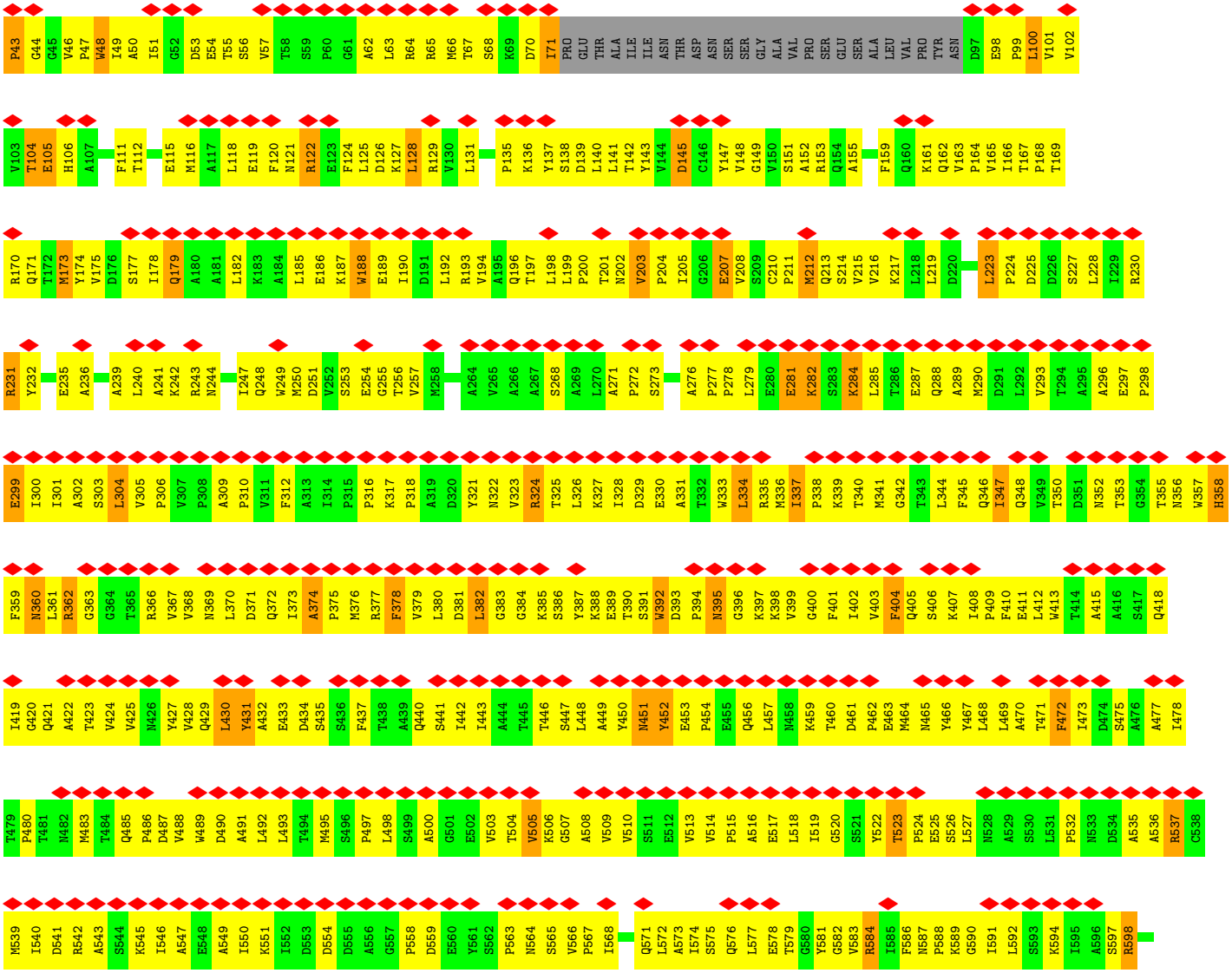
Chain	Residue	Modelled	Actual	Comment	Reference
A	944	SER	ALA	conflict	UNP Q91RA5
A	950	ARG	LYS	conflict	UNP Q91RA5
A	964	ASP	GLY	conflict	UNP Q91RA5
A	982	VAL	ILE	conflict	UNP Q91RA5
A	991	THR	ALA	conflict	UNP Q91RA5
A	992	ARG	LYS	conflict	UNP Q91RA5
A	1008	ILE	VAL	conflict	UNP Q91RA5
A	1031	ARG	GLN	conflict	UNP Q91RA5
A	1052	VAL	ILE	conflict	UNP Q91RA5
A	1083	THR	ALA	conflict	UNP Q91RA5
A	1119	VAL	ILE	conflict	UNP Q91RA5
A	1155	THR	ALA	conflict	UNP Q91RA5
A	1263	VAL	ILE	conflict	UNP Q91RA5
A	1274	ILE	LEU	conflict	UNP Q91RA5

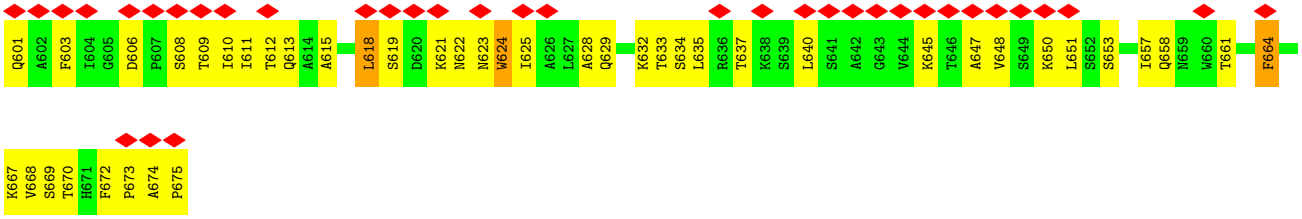
3 Residue-property plots (i)

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: Outer capsid protein mu-1

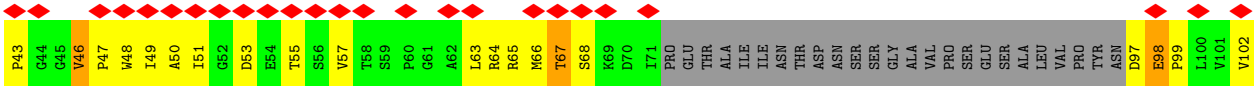
Chain K:

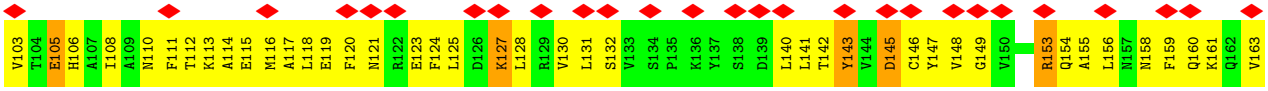


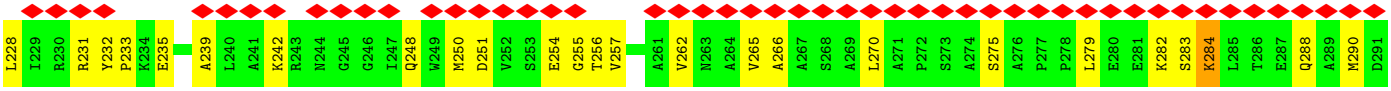


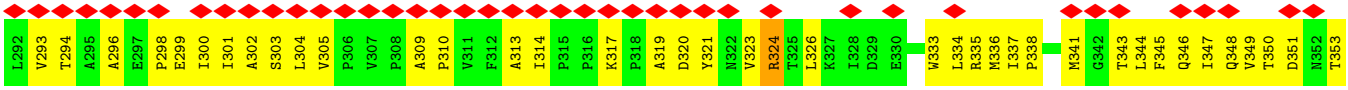
• Molecule 1: Outer capsid protein mu-1

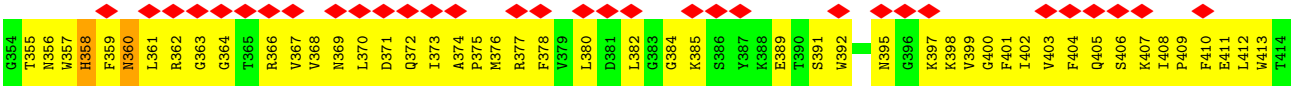
Chain F:

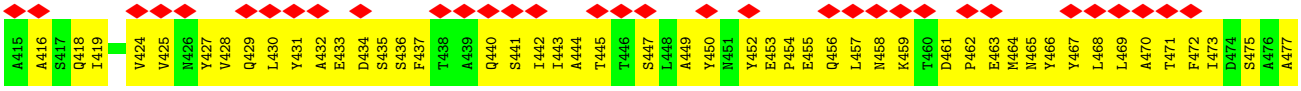


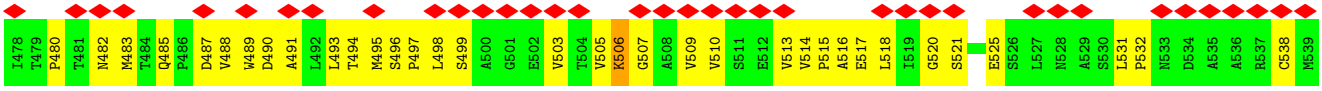


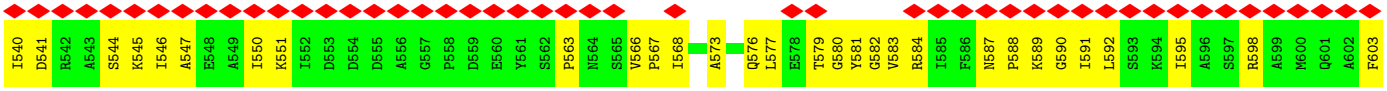


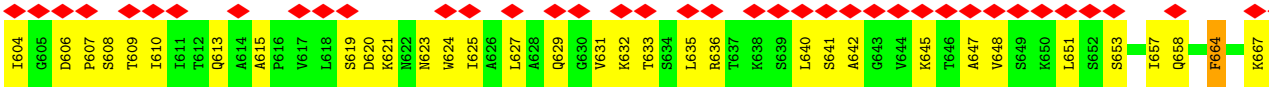






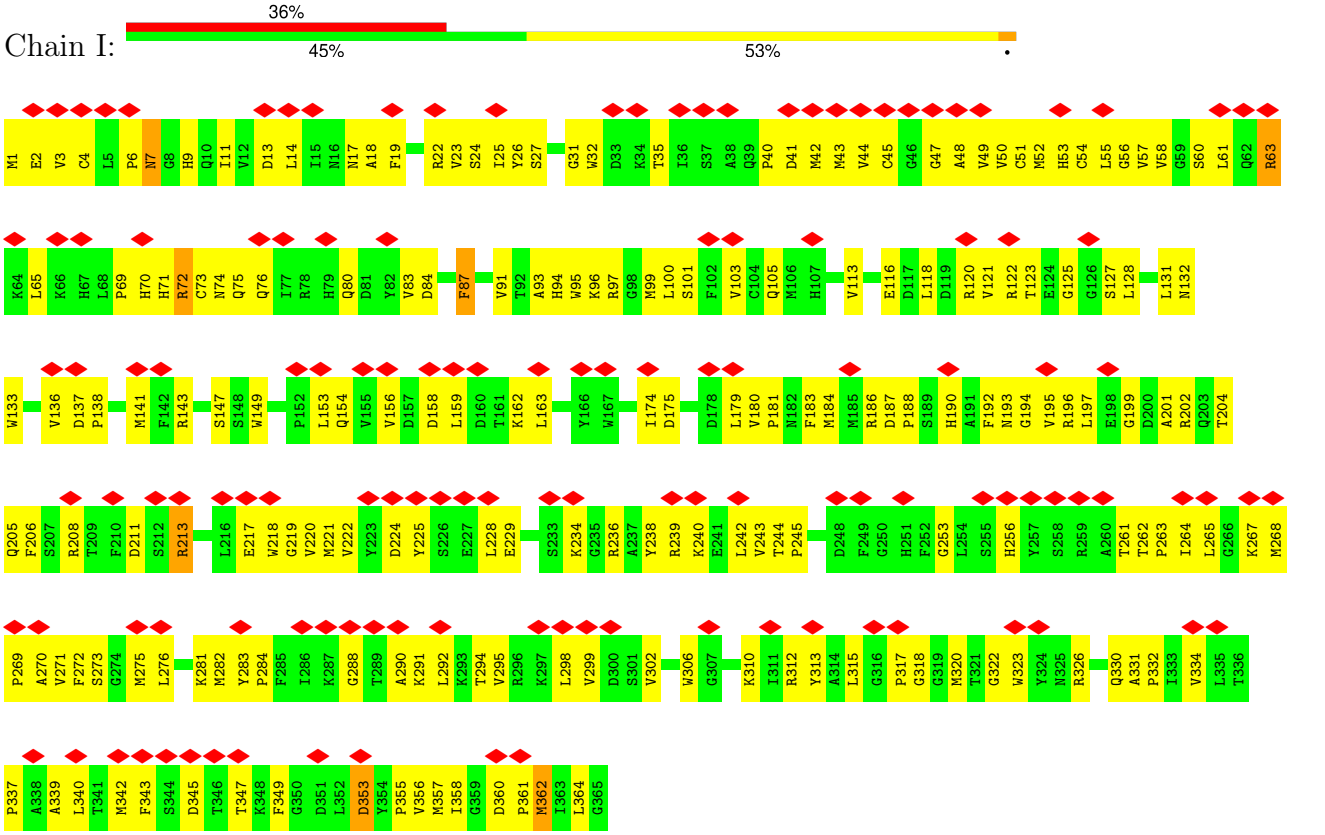






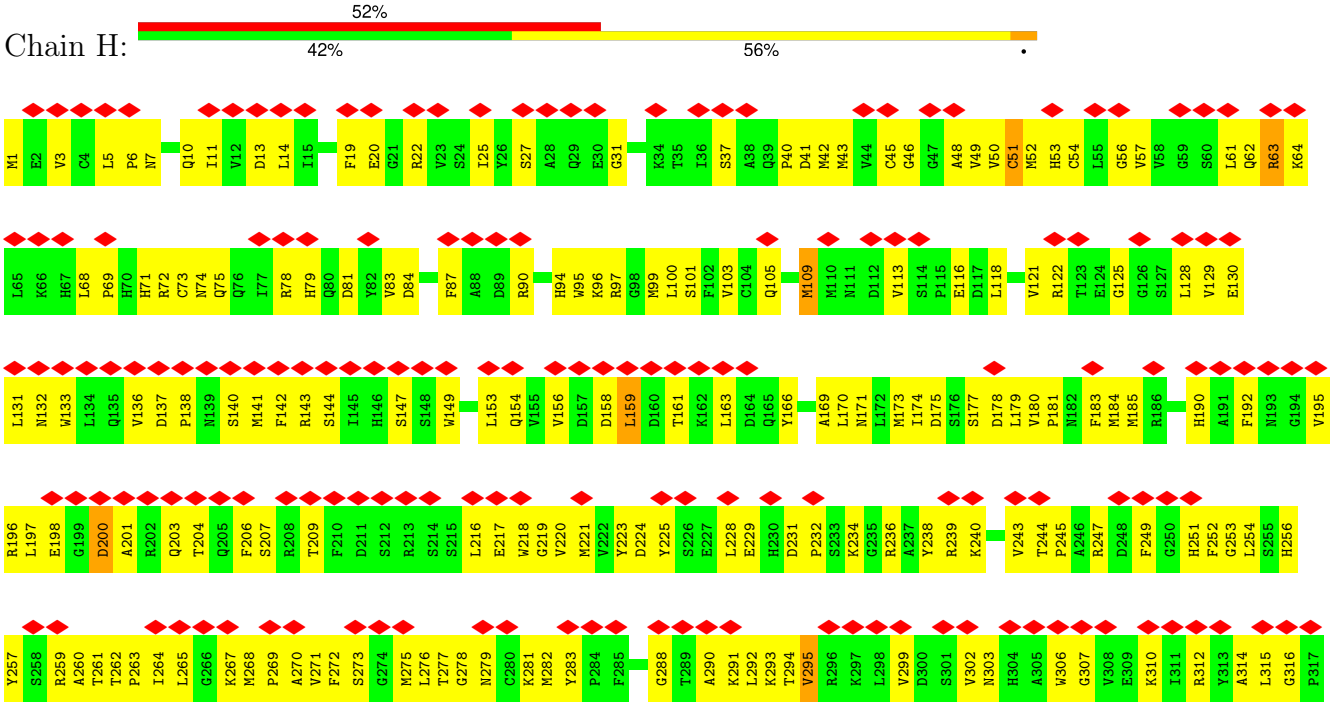
• Molecule 2: Outer capsid protein sigma-3

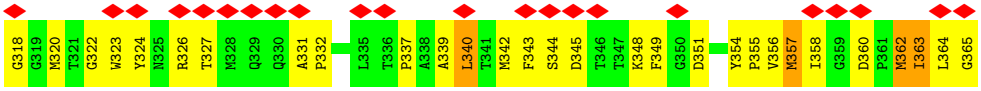
Chain I:



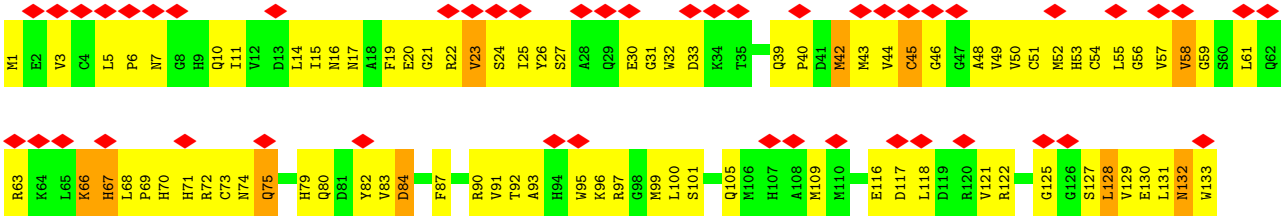
• Molecule 2: Outer capsid protein sigma-3

Chain H:

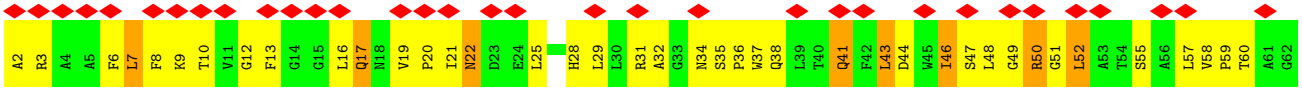


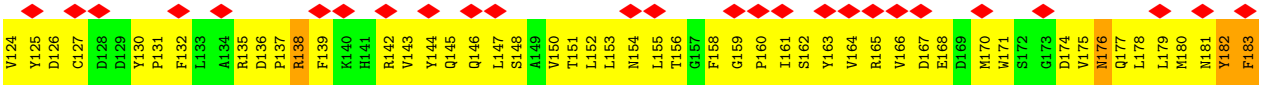


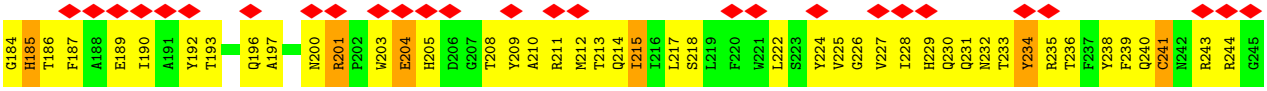
• Molecule 2: Outer capsid protein sigma-3

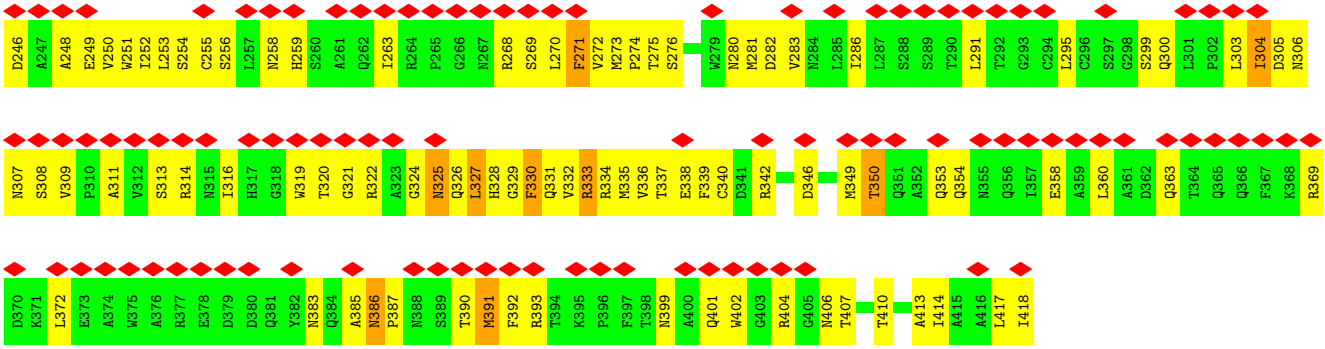


• Molecule 3: Inner capsid protein sigma-2



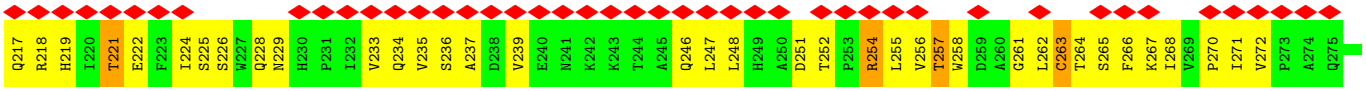


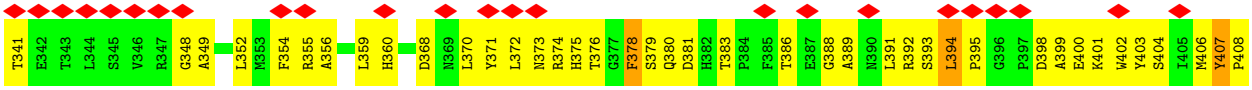


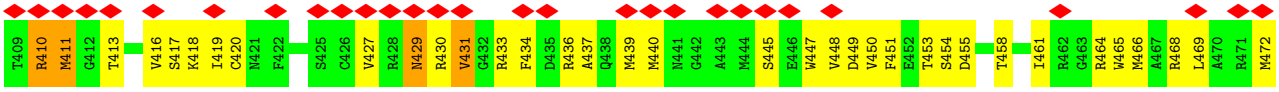


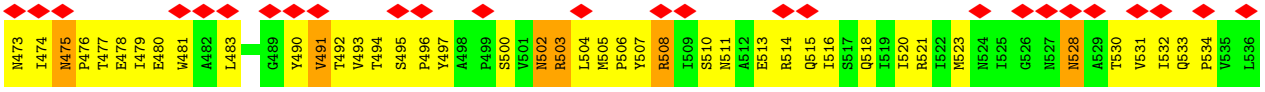
• Molecule 4: Inner capsid protein lambda-1

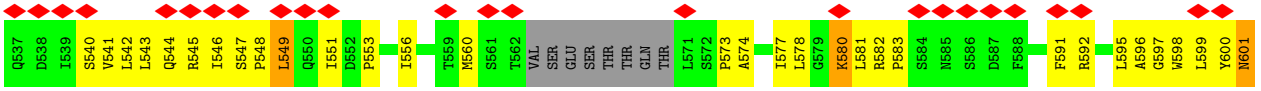
Chain C:

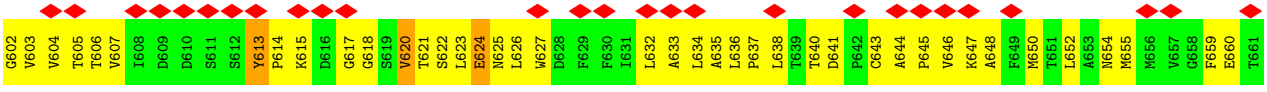


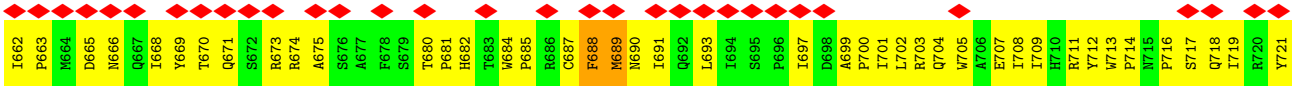


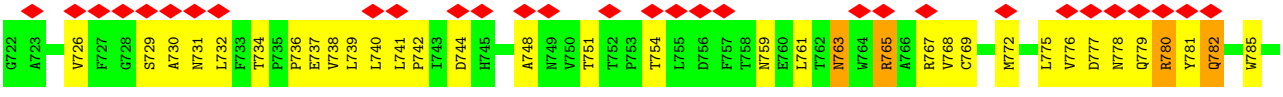


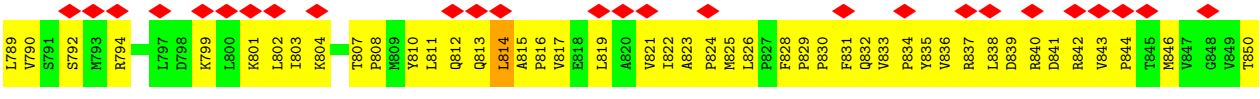


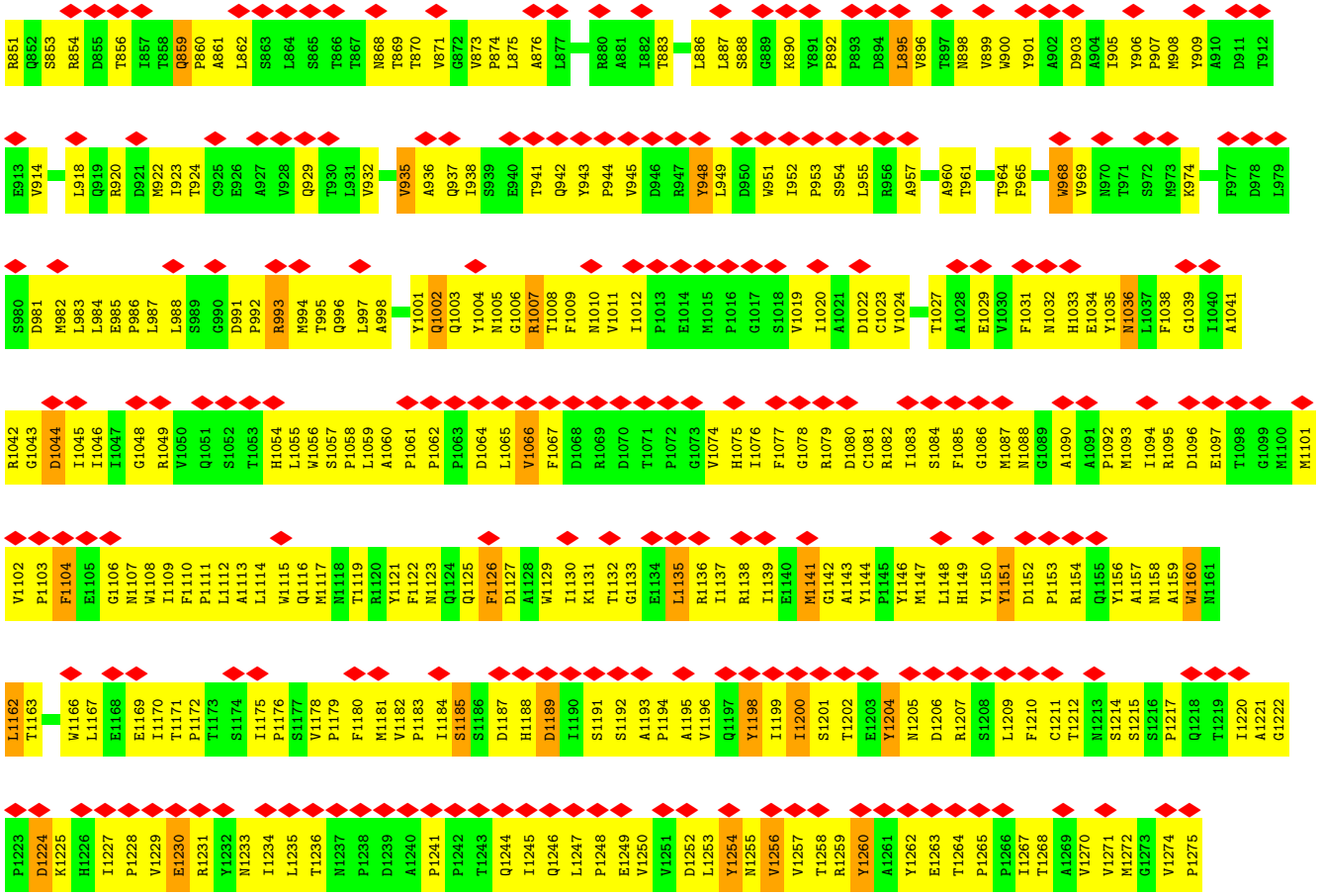






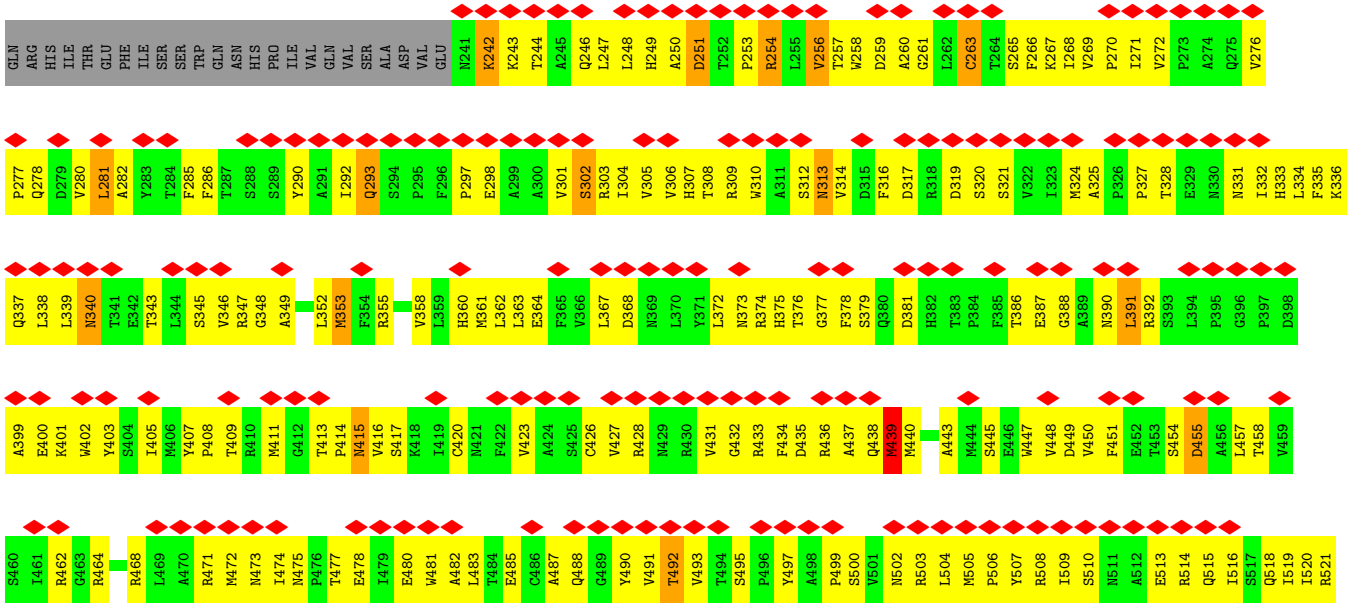


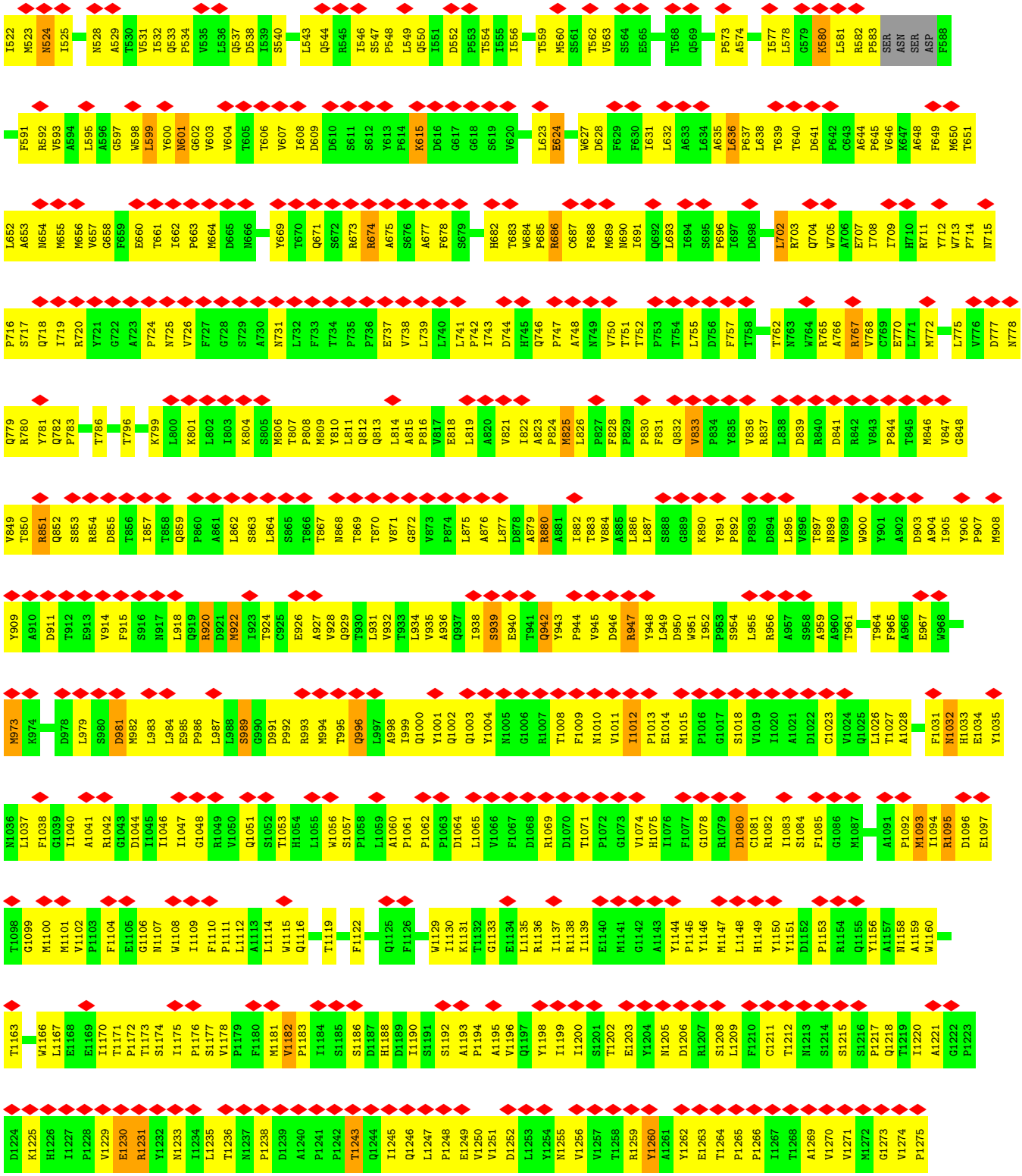




• Molecule 4: Inner capsid protein lambda-1

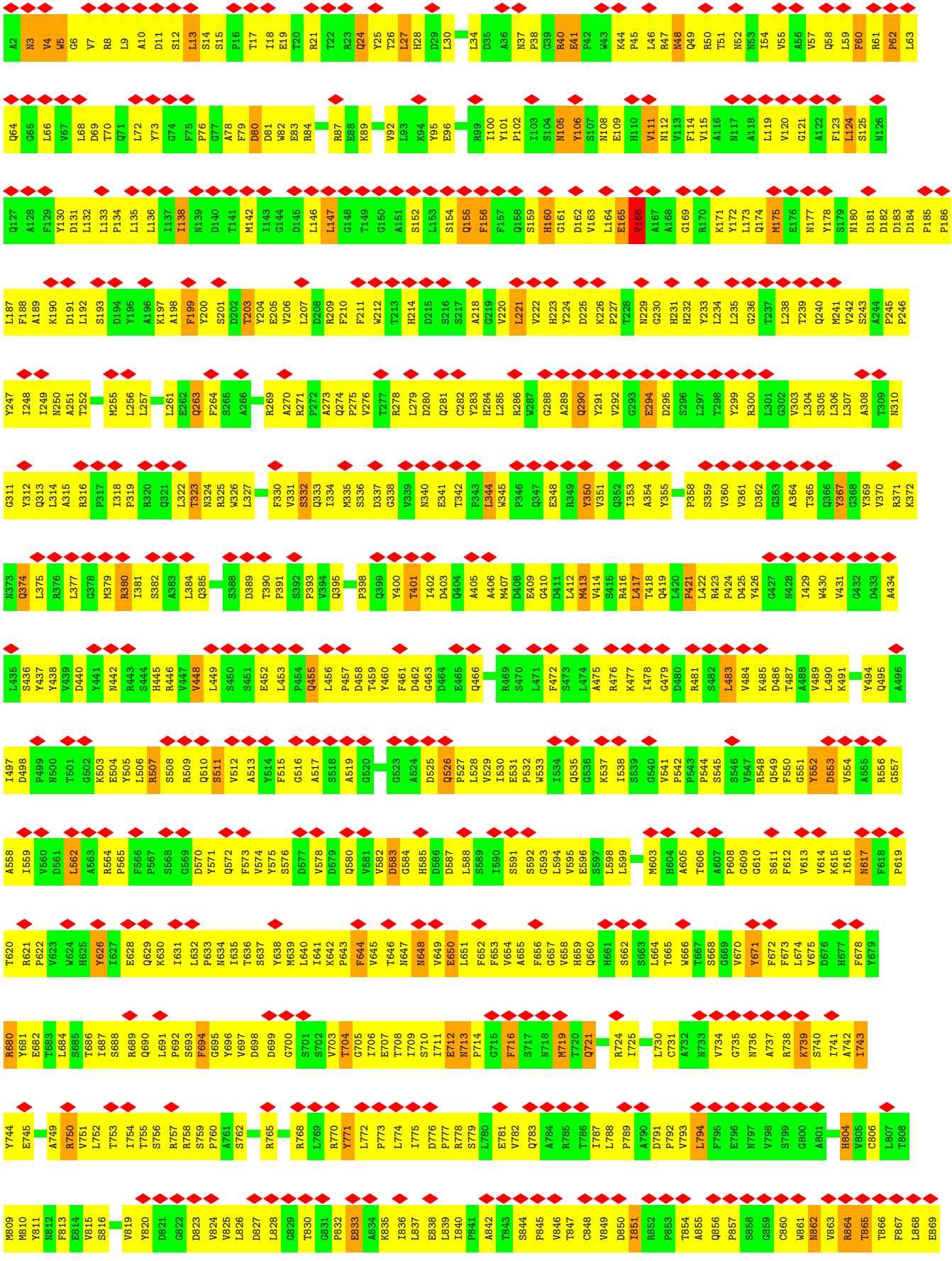
Chain B:

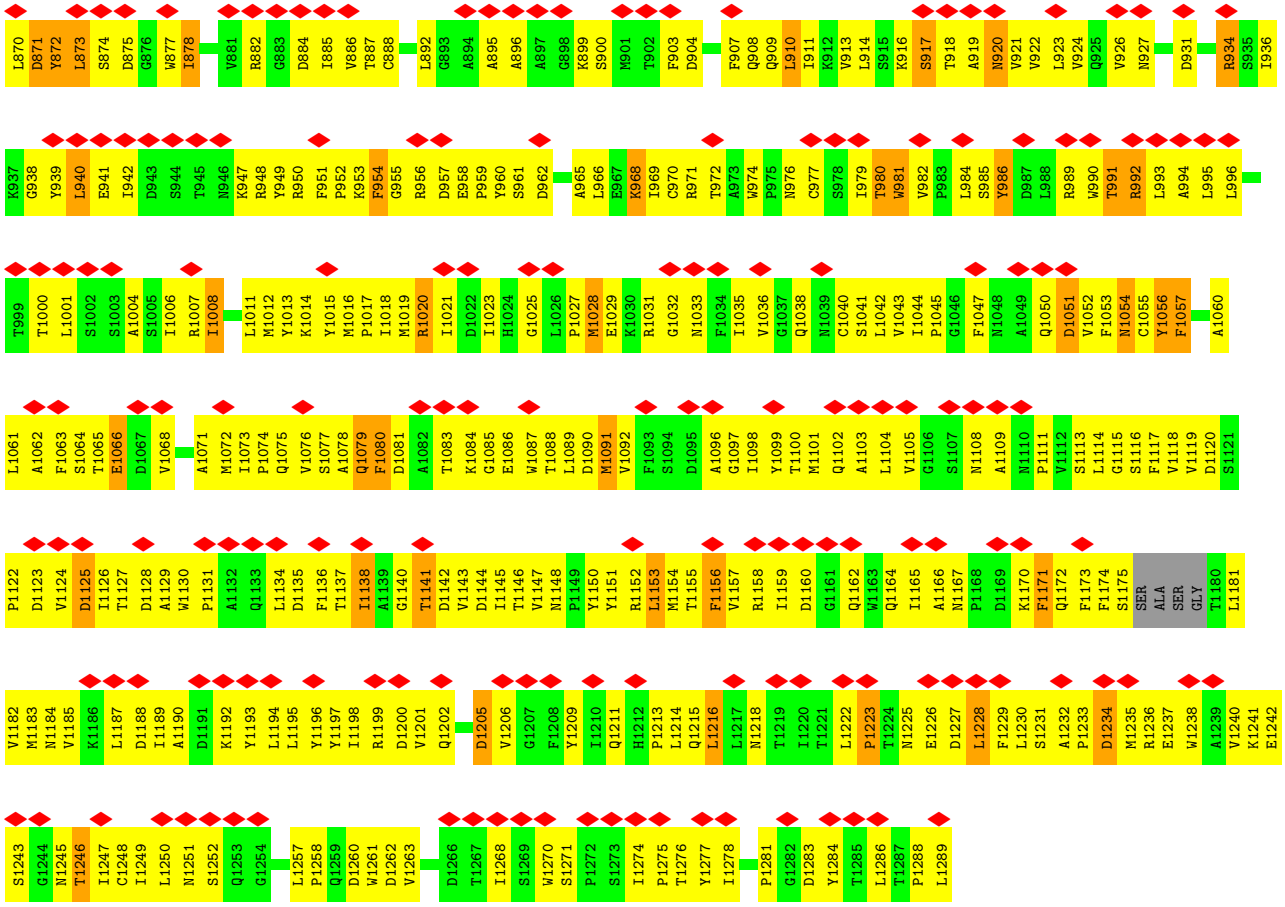




• Molecule 5: mRNA (guanine-N(7)-)-methyltransferase

Chain A:




L870 D871 Y872 L873 S874 D875 G876 W877 I878 V881 R882 G883 D884 I885 V886 T887 C888 L892 G893 A894 A895 A896 A897 G898 K899 S900 N901 T902 F903 D904 F907 Q908 Q909 L910 I911 K912 V913 L914 S915 K916 S917 T918 A919 N920 V921 V922 L923 V924 Q925 V926 N927 D931 R934 S935 I936
K937 G938 Y939 L940 E941 I942 D943 S944 T945 N946 K947 R948 Y949 R950 F951 P952 K953 F954 G955 R956 D957 E958 P959 Y960 S961 D962 A965 L966 E967 K968 I969 C970 R971 T972 A973 W974 P975 N976 C977 S978 I979 T980 W981 V982 P983 L984 S985 Y986 D987 I988 R989 W990 T991 R992 L993 A994 L995 L996
T999 T1000 L1001 S1002 S1003 A1004 S1005 I1006 R1007 I1008 L1011 M1012 Y1013 K1014 Y1015 M1016 P1017 I1018 M1019 R1020 I1021 D1022 I1023 H1024 G1025 L1026 P1027 M1028 E1029 K1030 R1031 G1032 N1033 F1034 I1035 V1036 G1037 Q1038 N1039 C1040 S1041 L1042 V1043 I1044 P1045 G1046 F1047 N1048 A1049 Q1050 D1051 V1052 F1053 N1054 C1055 Y1056 F1057 A1060
L1061 A1062 F1063 S1064 T1065 E1066 D1067 V1068 A1071 M1072 I1073 P1074 Q1075 V1076 S1077 A1078 Q1079 F1080 D1081 A1082 T1083 K1084 G1085 E1086 W1087 T1088 L1089 D1090 M1091 V1092 F1093 S1094 D1095 A1096 G1097 I1098 Y1099 T1100 M1101 Q1102 A1103 L1104 V1105 G1106 S1107 N1108 A1109 N1110 P1111 V1112 S1113 L1114 G1115 S1116 F1117 V1118 V1119 D1120 S1121
P1122 D1123 V1124 D1125 I1126 T1127 D1128 A1129 W1130 P1131 A1132 Q1133 L1134 D1135 F1136 T1137 I1138 A1139 G1140 T1141 D1142 V1143 D1144 I1145 T1146 V1147 N1148 P1149 Y1150 Y1151 R1152 L1153 M1154 T1155 F1156 V1157 R1158 I1159 D1160 G1161 Q1162 W1163 Q1164 I1165 A1166 N1167 P1168 D1169 K1170 F1171 Q1172 F1173 F1174 S1175 SER ALA SER GLY T1180 L1181
V1182 M1183 N1184 V1185 K1186 L1187 D1188 I1189 A1190 D1191 K1192 Y1193 L1194 L1195 Y1196 Y1197 I1198 R1199 D1200 V1201 Q1202 D1205 V1206 G1207 F1208 Y1209 I1210 Q1211 H1212 P1213 L1214 Q1215 L1216 L1217 N1218 T1219 I1220 T1221 L1222 P1223 T1224 N1225 E1226 D1227 L1228 F1229 L1230 S1231 A1232 P1233 D1234 M1235 R1236 E1237 W1238 A1239 V1240 K1241 E1242
S1243 G1244 N1245 T1246 I1247 C1248 I1249 L1250 N1251 S1252 Q1253 G1254 L1257 P1258 Q1259 D1260 W1261 D1262 V1263 D1266 T1267 I1268 S1269 W1270 S1271 P1272 S1273 I1274 P1275 T1276 Y1277 I1278 P1281 G1282 D1283 Y1284 T1285 L1286 T1287 P1288 L1289

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, C5	Depositor
Number of subtomograms used	625	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 ($e^-/Å^2$)	2	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	8.955	Depositor
Minimum map value	-5.801	Depositor
Average map value	0.054	Depositor
Map value standard deviation	0.430	Depositor
Recommended contour level	2.7	Depositor
Map size (Å)	648.0, 648.0, 648.0	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.8, 1.8, 1.8	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	F	0.63	0/4737	0.68	0/6466
1	K	0.62	0/4737	0.69	0/6466
2	G	0.62	0/2957	0.69	0/4005
2	H	0.61	0/2957	0.68	0/4005
2	I	0.61	0/2957	0.68	0/4005
3	E	0.59	0/3398	0.68	0/4626
4	B	0.60	0/8363	0.69	0/11454
4	C	0.61	0/8537	0.68	0/11693
5	A	0.61	0/10384	0.68	0/14170
All	All	0.61	0/49027	0.68	0/66890

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	F	4641	0	4668	603	0
1	K	4641	0	4655	948	0
2	G	2885	0	2806	465	0
2	H	2885	0	2813	297	0
2	I	2885	0	2813	254	0
3	E	3313	0	3214	570	0
4	B	8143	0	8056	1003	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	8311	0	8205	1030	0
5	A	10126	0	9903	1406	0
All	All	47830	0	47133	5917	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 62.

The worst 5 of 5917 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:K:506:LYS:HG3	2:G:313:TYR:CG	1.26	1.65
1:K:584:ARG:HB2	2:G:70:HIS:CD2	1.27	1.63
3:E:258:ASN:HB2	4:C:854:ARG:CZ	1.30	1.61
1:K:504:THR:HG22	2:G:317:PRO:CG	1.13	1.60
4:C:394:LEU:HD22	4:C:395:PRO:CD	1.27	1.58

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	F	604/633 (95%)	593 (98%)	11 (2%)	0	100	100
1	K	604/633 (95%)	598 (99%)	5 (1%)	1 (0%)	44	78
2	G	363/365 (100%)	359 (99%)	3 (1%)	1 (0%)	37	73
2	H	363/365 (100%)	357 (98%)	6 (2%)	0	100	100
2	I	363/365 (100%)	360 (99%)	3 (1%)	0	100	100
3	E	415/417 (100%)	407 (98%)	7 (2%)	1 (0%)	44	78
4	B	1027/1059 (97%)	1014 (99%)	12 (1%)	1 (0%)	48	83
4	C	1047/1059 (99%)	1026 (98%)	21 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	A	1280/1288 (99%)	1244 (97%)	28 (2%)	8 (1%)	22 60
All	All	6066/6184 (98%)	5958 (98%)	96 (2%)	12 (0%)	45 78

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	A	62	PRO
1	K	374	ALA
3	E	52	LEU
4	B	439	MET
5	A	1109	ALA

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	F	513/534 (96%)	493 (96%)	20 (4%)	27 48
1	K	513/534 (96%)	462 (90%)	51 (10%)	6 21
2	G	317/317 (100%)	288 (91%)	29 (9%)	7 24
2	H	317/317 (100%)	302 (95%)	15 (5%)	22 44
2	I	317/317 (100%)	302 (95%)	15 (5%)	22 44
3	E	352/352 (100%)	319 (91%)	33 (9%)	7 23
4	B	911/938 (97%)	838 (92%)	73 (8%)	10 29
4	C	930/938 (99%)	846 (91%)	84 (9%)	8 24
5	A	1118/1120 (100%)	986 (88%)	132 (12%)	4 16
All	All	5288/5367 (98%)	4836 (92%)	452 (8%)	11 27

5 of 452 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	C	1141	MET
5	A	1156	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	B	833	VAL
5	A	1123	ASP
5	A	771	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 149 such sidechains are listed below:

Mol	Chain	Res	Type
5	A	71	GLN
5	A	908	GLN
5	A	160	HIS
5	A	385	GLN
4	C	229	ASN

5.3.3 RNA

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

There are no ligands in this entry.

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.

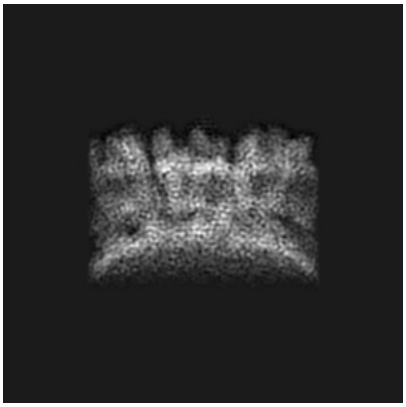
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-22166. These allow visual inspection of the internal detail of the map and identification of artifacts.

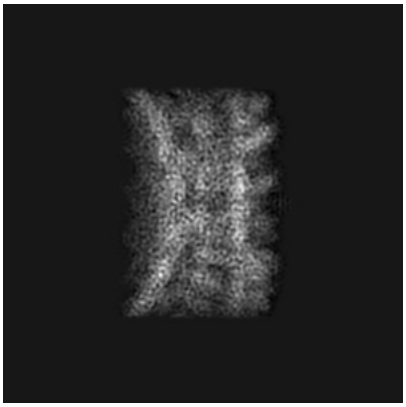
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

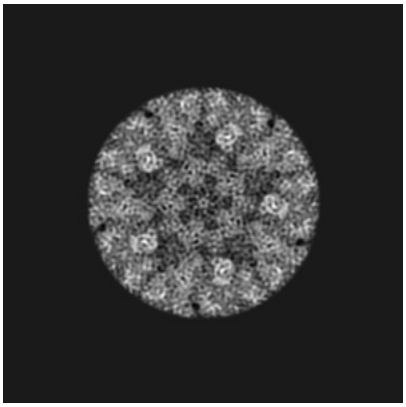
6.1.1 Primary map



X



Y

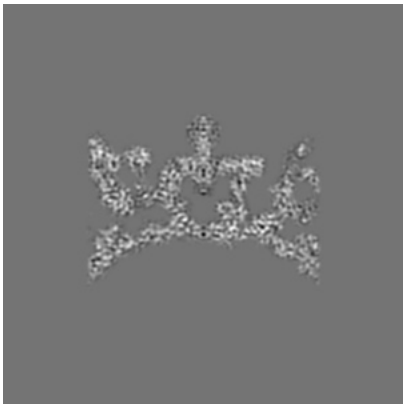


Z

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

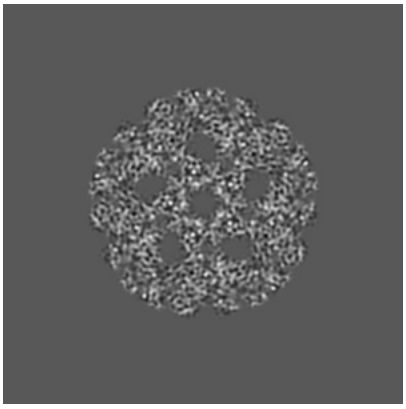
6.2 Central slices

6.2.1 Primary map



X Index: 180

Y Index: 180



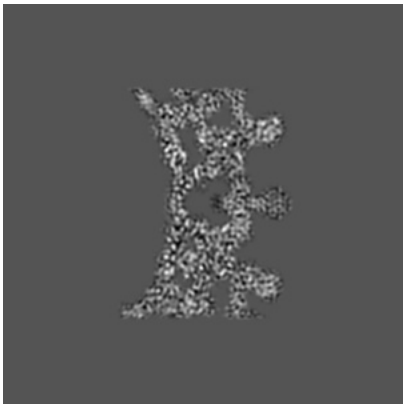
Z Index: 180

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

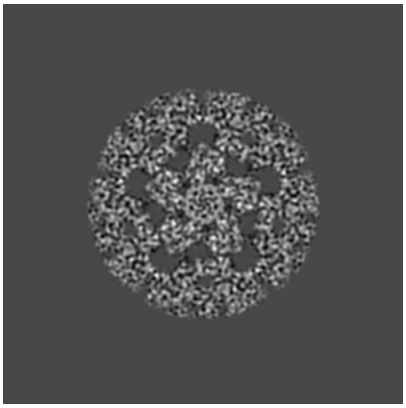
6.3 Largest variance slices (i)

6.3.1 Primary map

X Index: 195



Y Index: 176

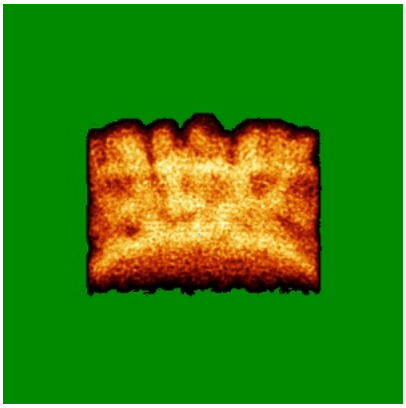


Z Index: 213

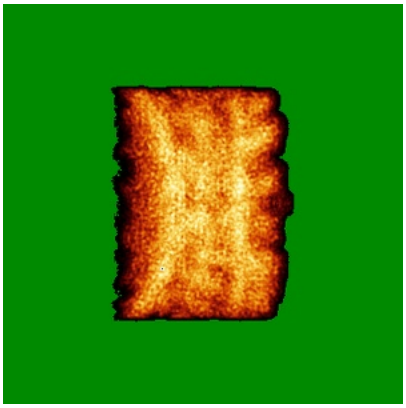
The images above show the largest variance slices of the map in three orthogonal directions.

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

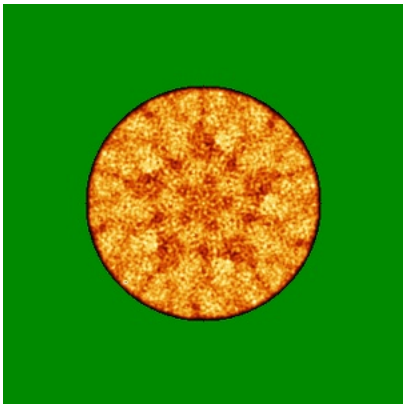
6.4.1 Primary map



X



Y

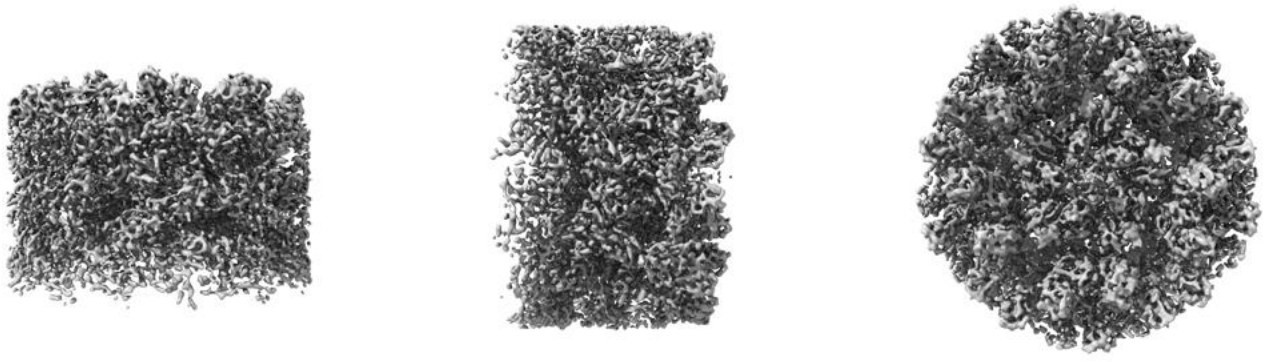


Z

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

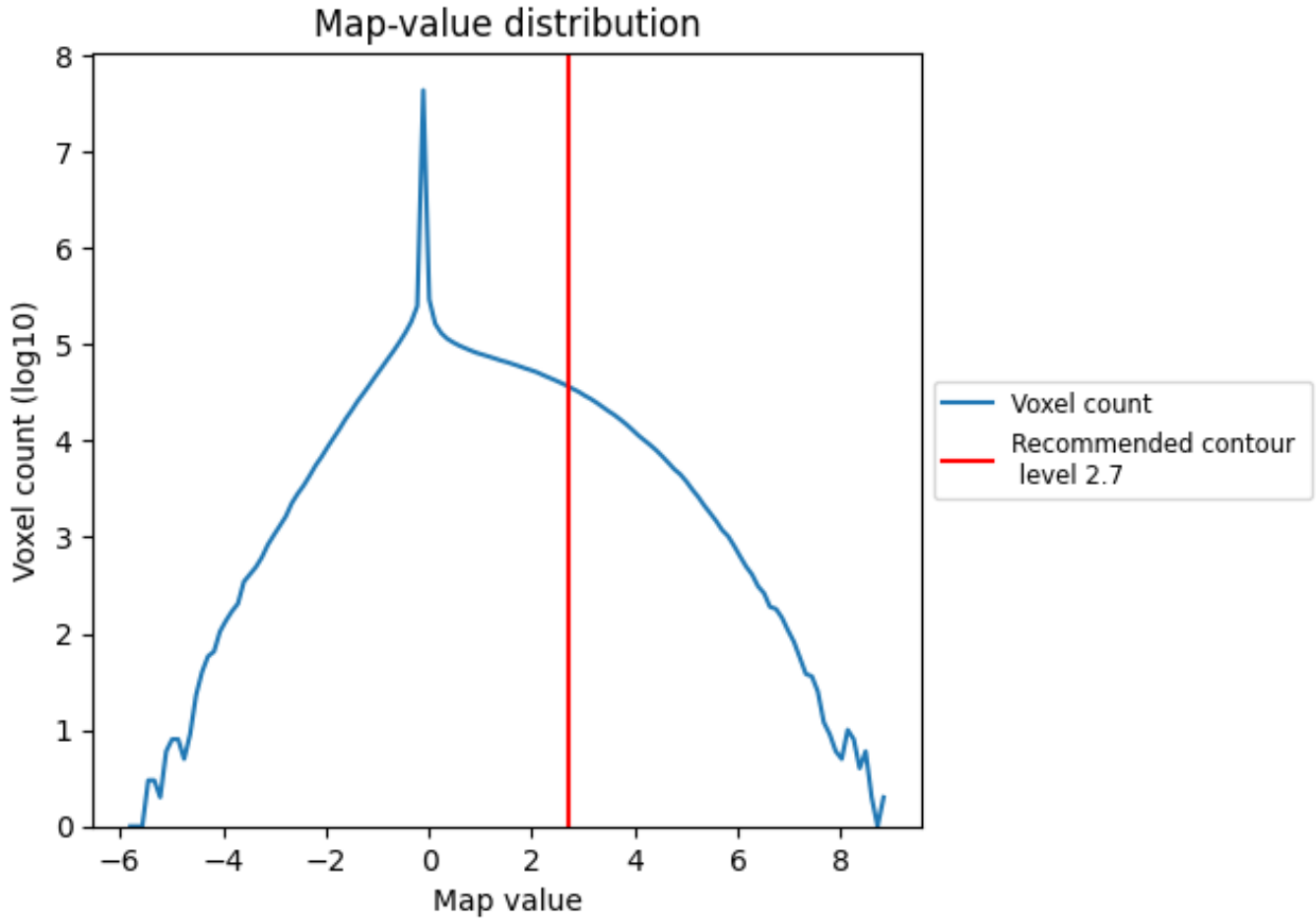
6.6 Mask visualisation (i)

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

7 Map analysis

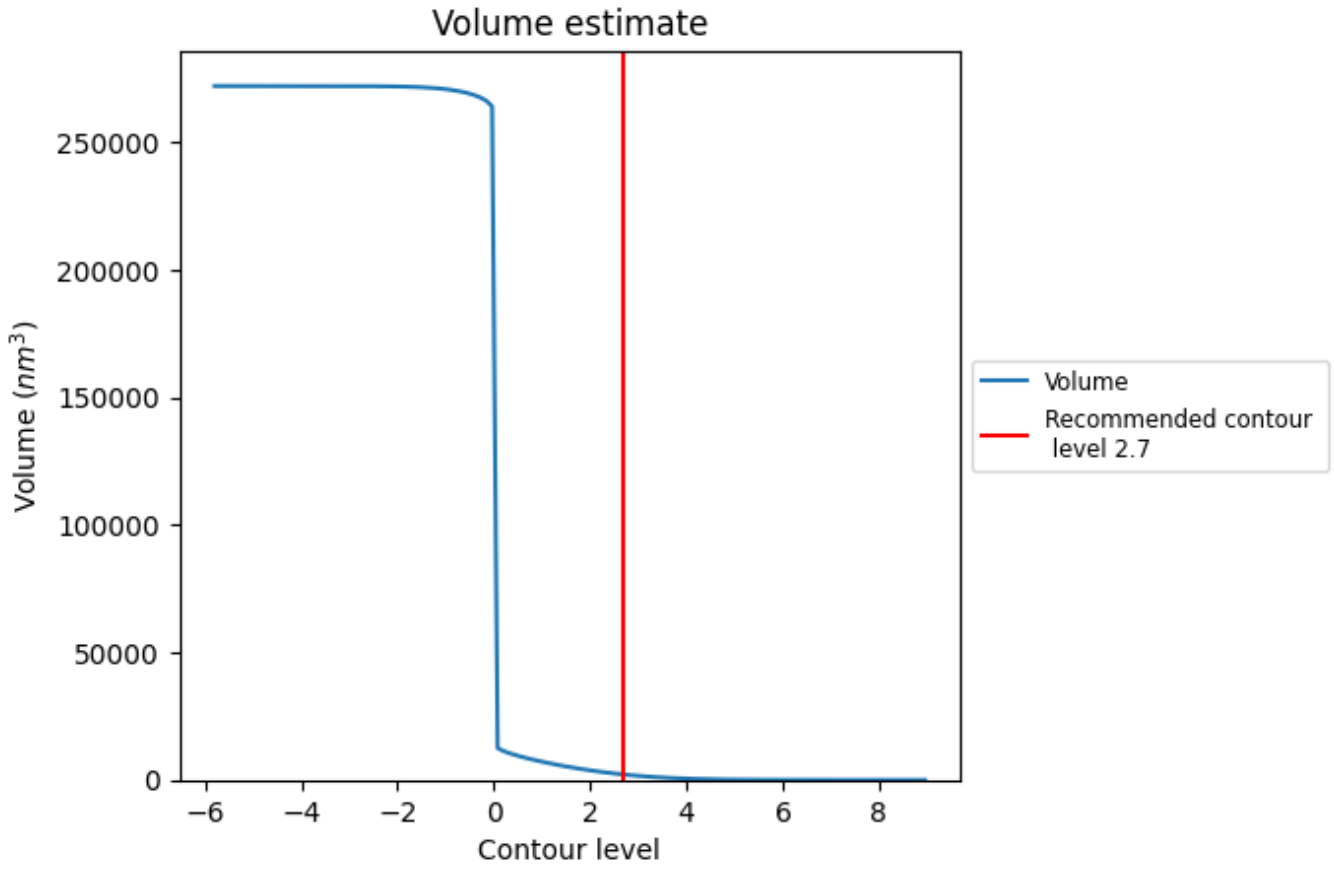
This section contains the results of statistical analysis of the map.

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.

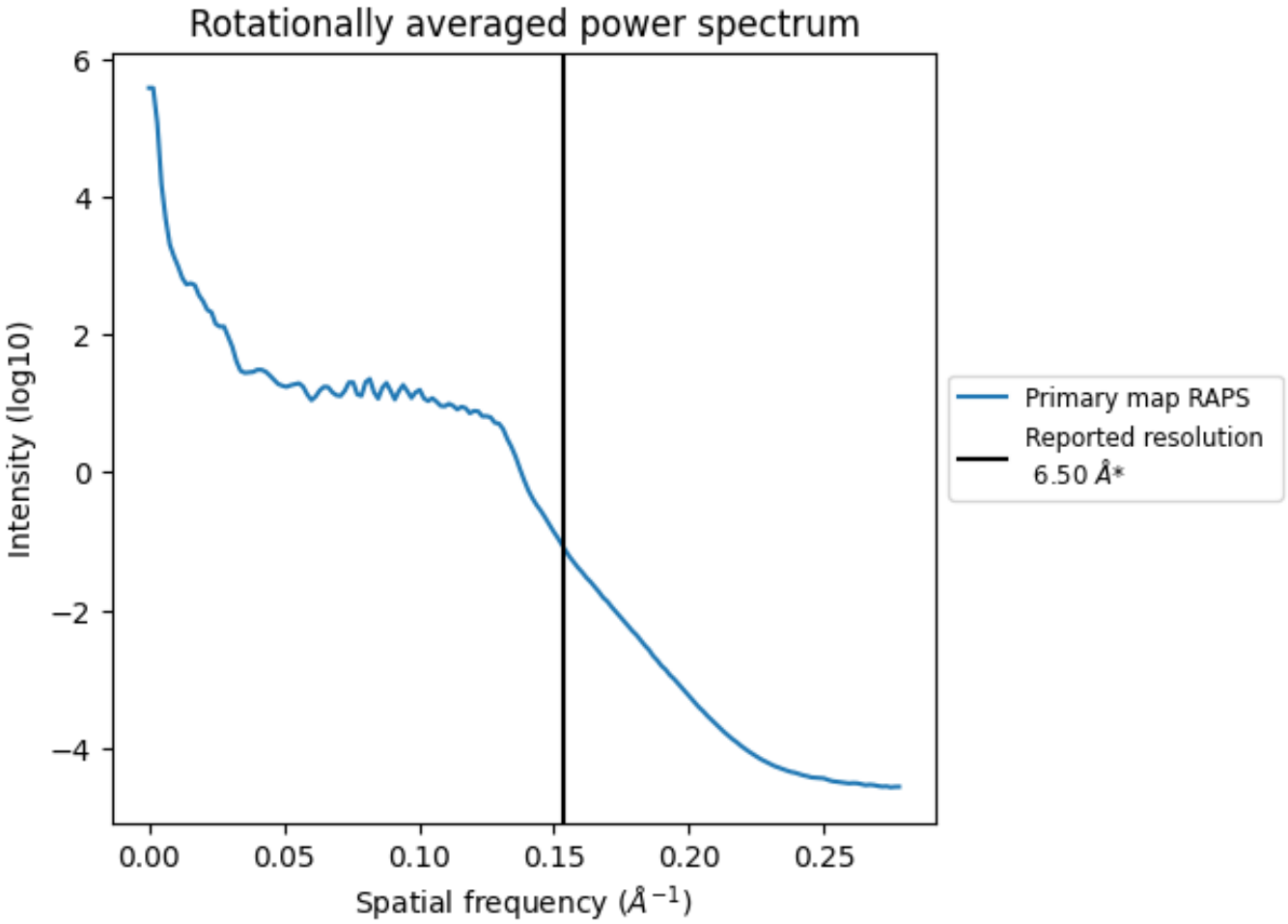
7.2 Volume estimate (i)



The volume at the recommended contour level is 2134 nm^3; this corresponds to an approximate mass of 1928 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 0.154 Å^{-1}

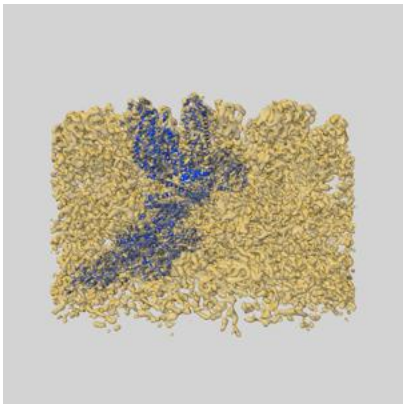
8 Fourier-Shell correlation (i)

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

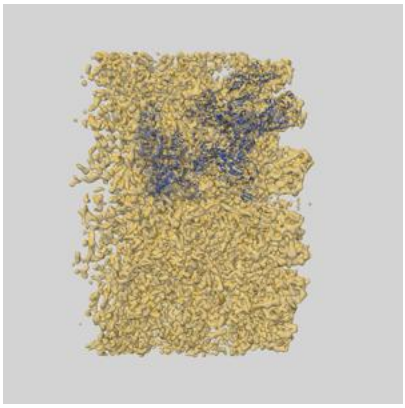
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-22166 and PDB model 6XF8. Per-residue inclusion information can be found in section 3 on page 7.

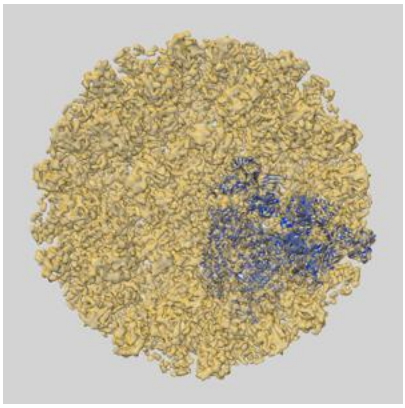
9.1 Map-model overlay



X



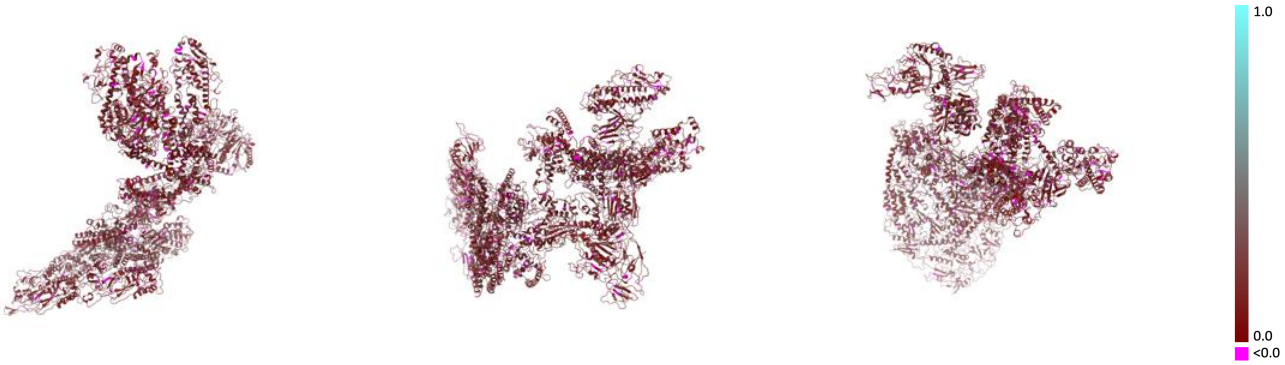
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 2.7 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.

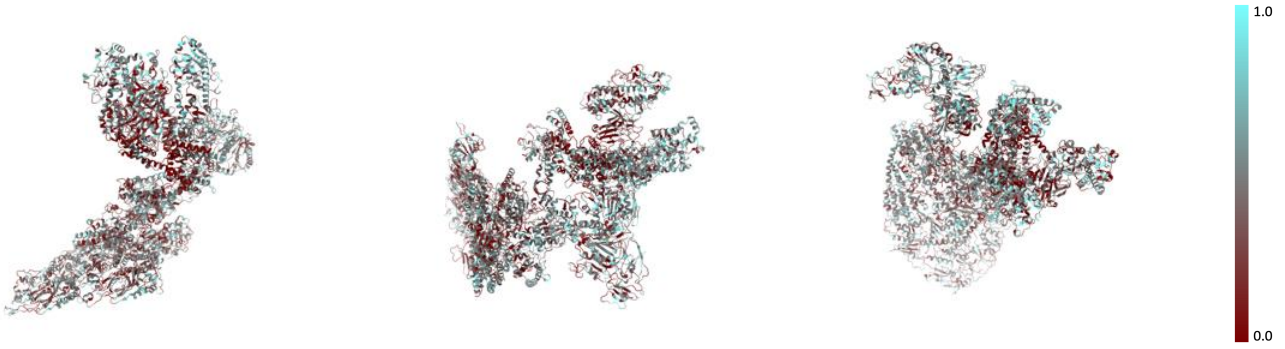
9.2 Q-score mapped to coordinate model



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.

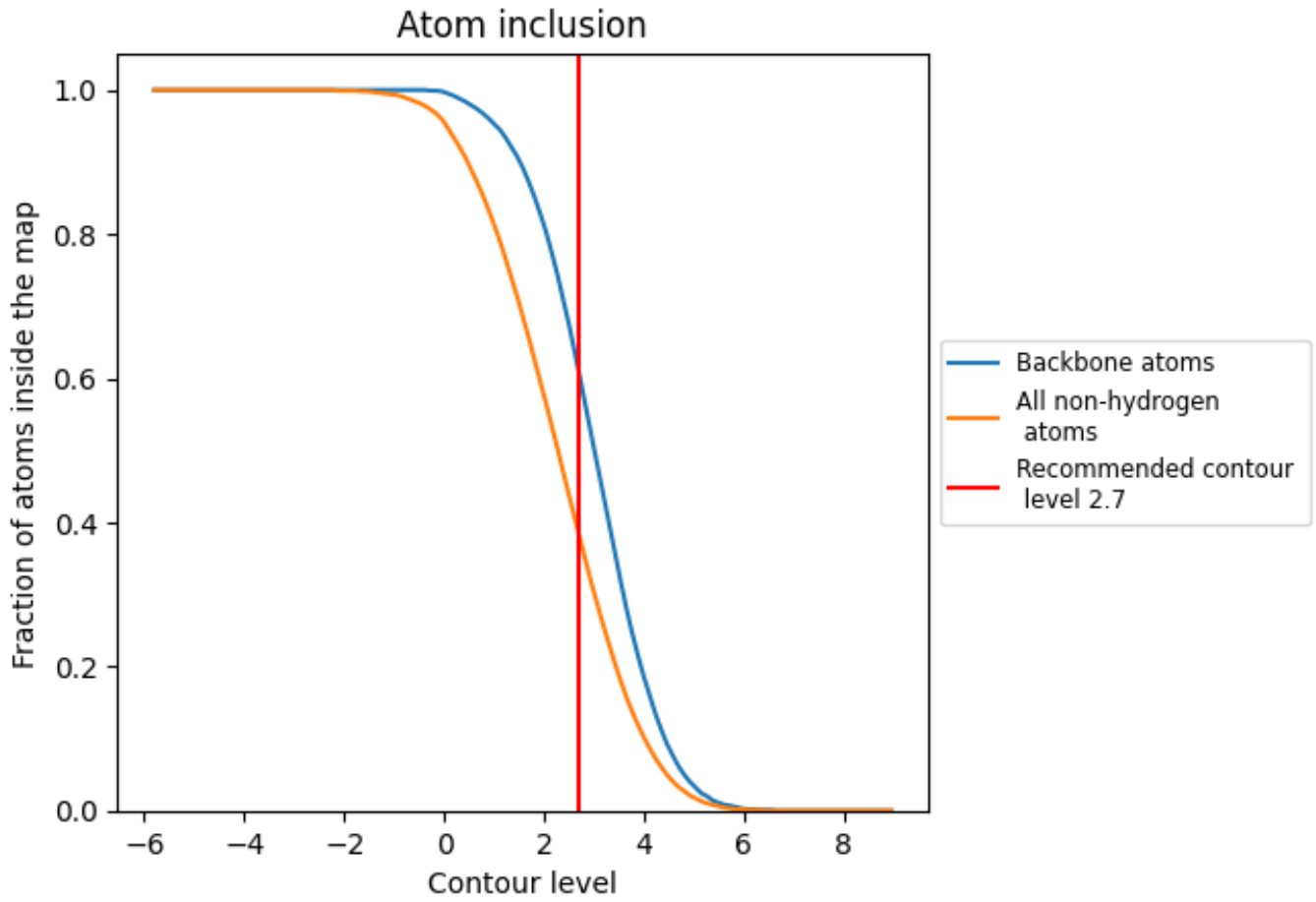
9.3 Atom inclusion mapped to coordinate model



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 (2.7).

9.4 Atom inclusion (i)



At the recommended contour level, 61% of all backbone atoms, 38% 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 (2.7) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.3830	0.1800
A	0.4180	0.1900
B	0.3490	0.1770
C	0.3920	0.1840
E	0.3840	0.1770
F	0.3240	0.1690
G	0.4760	0.1890
H	0.3860	0.1640
I	0.4830	0.1750
K	0.2910	0.1760