wwPDB EM Validation Summary Report (i)

Nov 19, 2022 – 06:25 pm GMT

PDB ID	:	5OEJ
EMDB ID	:	EMD-3790
Title	:	Structure of Tra1 subunit within the chromatin modifying complex SAGA
Authors	:	Sharov, G.; Voltz, K.; Durand, A.; Kolesnikova, O.; Papai, G.; Myasnikov, A.G.; Dejaegere, A.; Ben-Shem, A.; Schultz, P.
Deposited on	:	2017-07-07
Resolution	:	5.70 Å(reported)
Based on initial model	:	4JSN

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:

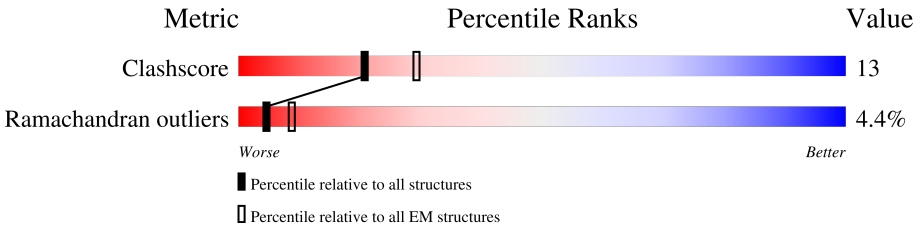
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 5.70 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

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	B	3825	19% 60% 14% 26%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 14023 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 Tra1 subunit within the chromatin modifying complex SAGA.

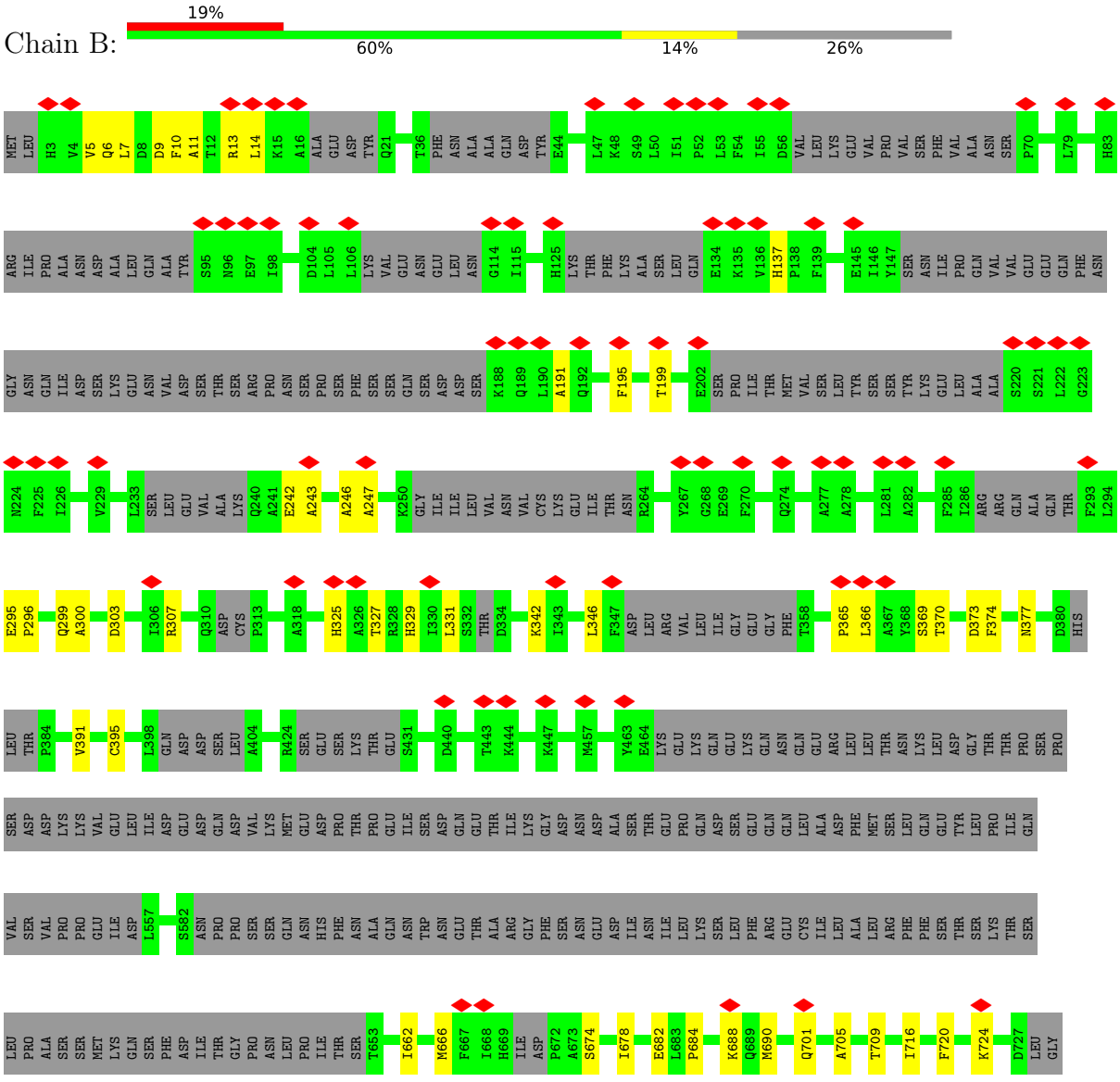
Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	B	2825	14023	8373	2825	2825	0	0

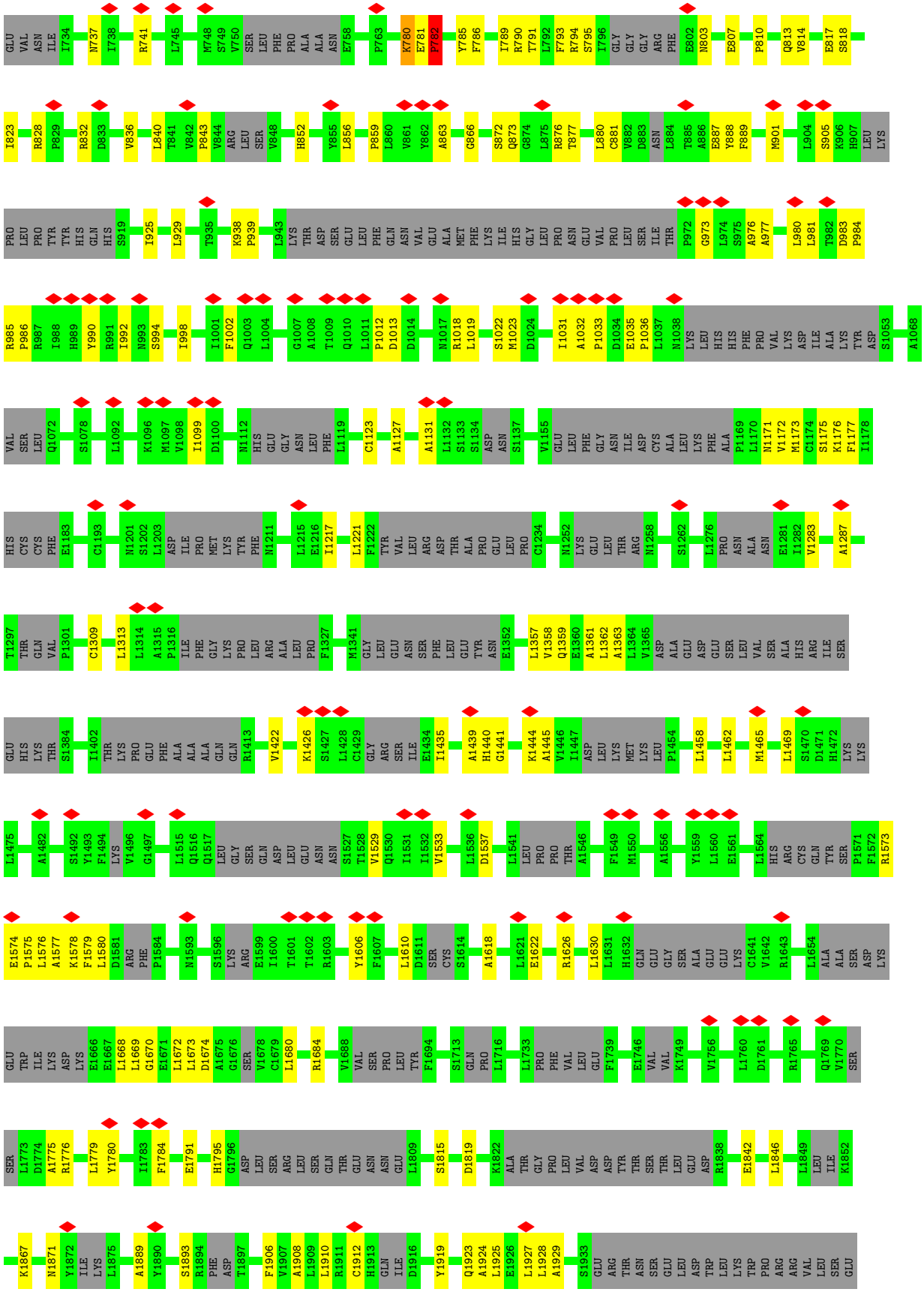
3 Residue-property plots (i)

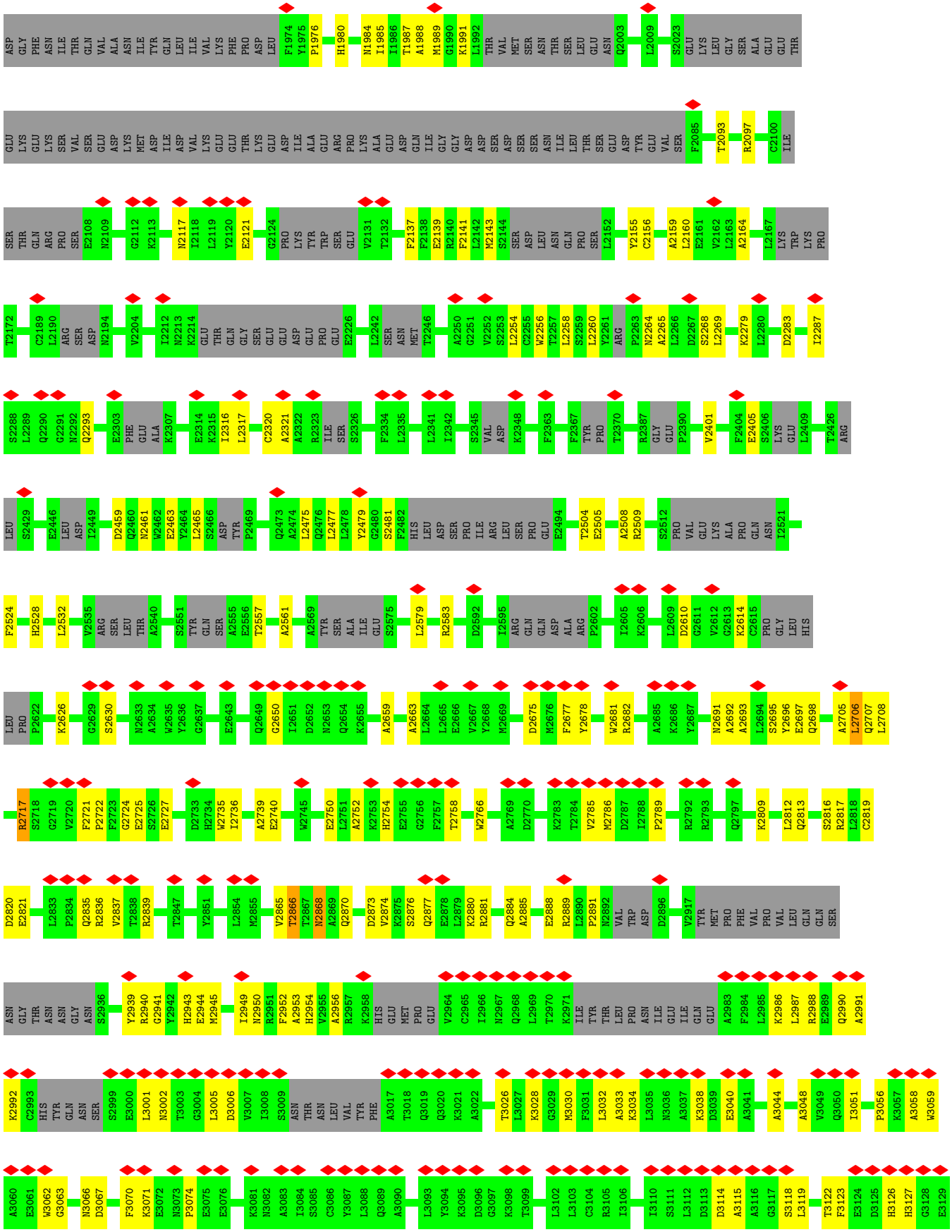
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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.

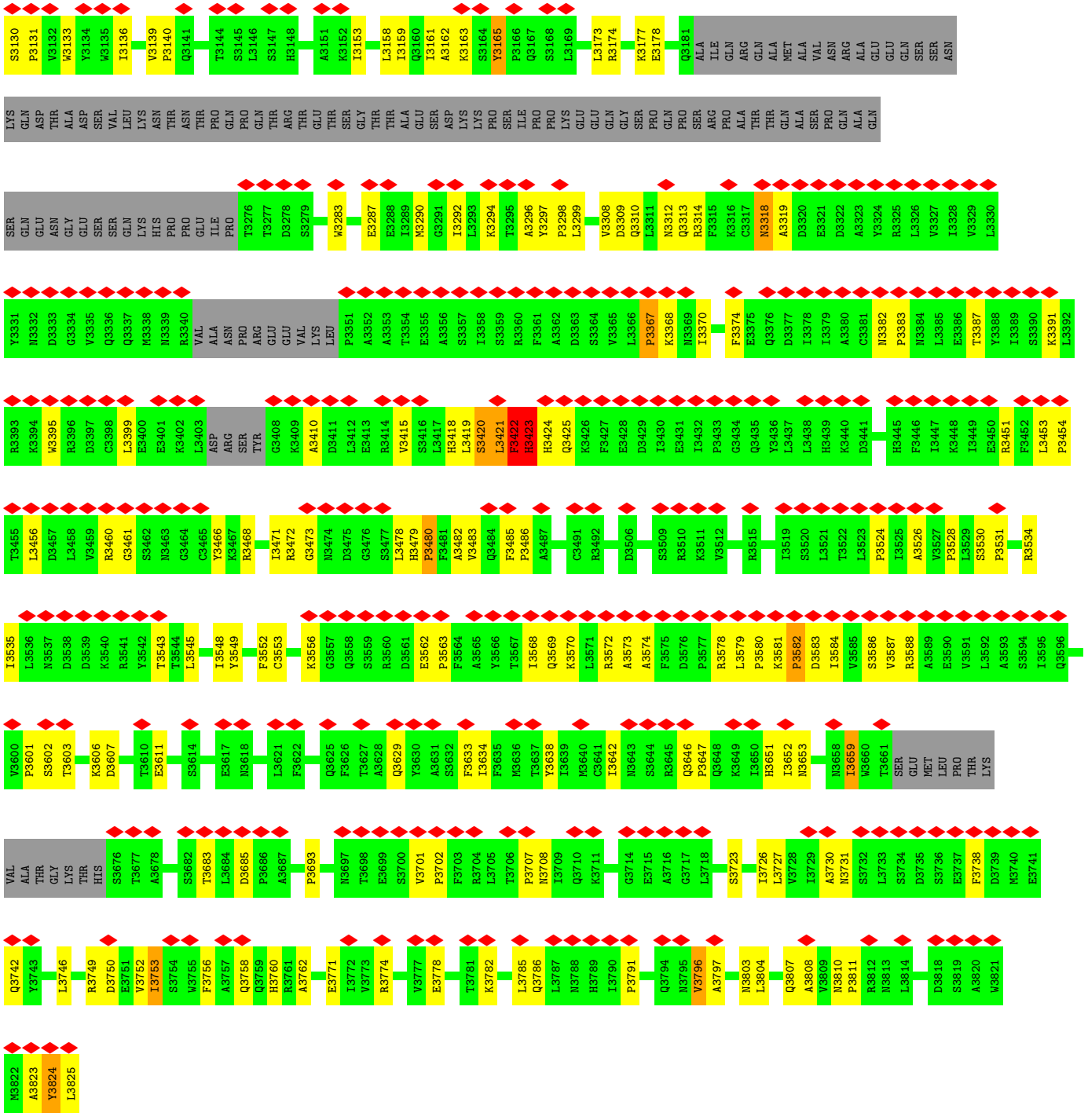
- Molecule 1: Tra1 subunit within the chromatin modifying complex SAGA

Chain B:




GLU VAL ASN ILE I734 N737 I738 R741 L745 M748 S749 V750 SER LEU PHE PRO ALA ALA ASN E768 P763 K780 E781 P782 Y785 F786 I789 R790 T791 L792 F793 R794 S795 I796 GLY GLY GLY ARG PHE E802 N803 E807 P810 Q813 V814 E817 S818
I823 R828 P829 R832 D833 V836 L840 T841 V842 P843 V844 ARG LEU SER V848 H852 Y855 L856 P859 L860 V861 Y862 A863 G866 S872 Q873 G874 L875 R876 T877 L880 C881 V882 D883 ASN L884 T885 A886 E887 Y888 F889 M901 L904 S905 K906 H907 LEU LYS
PRO LEU PRO TYR TYR HIS GLN HIS S919 I925 L929 T935 K938 P939 L943 LYS THR ASP SER GLU LEU PHE GLN ASN VAL GLU ALA MET PHE LYS ILE HIS GLY LEU PRO ASN GLU VAL PRO LEU SER ILE THR P972 G973 L974 S975 A976 A977 L980 L981 T982 D983 P984
R985 P986 R987 I988 H989 Y990 R991 I992 N993 S994 I998 I1001 F1002 Q1003 L1004 G1007 A1008 T1009 Q1010 L1011 P1012 D1013 D1014 N1017 R1018 L1019 S1022 M1023 D1024 I1031 A1032 P1033 D1034 E1035 P1036 L1037 N1038 LYS LEU HIS HIS PHE PRO VAL LYS ASP ILE ALA LYS TYR ASP S1053 A1068
VAL SER LEU Q1072 S1078 L1092 K1096 M1097 V1098 I1099 D1100 N1112 HIS GLU GLY ASN LEU PHE L1119 C1123 A1127 A1131 L1132 S1133 S1134 ASP ASN S1137 V1155 GLU LEU PHE GLY ASN ILE ASP CYS ALA LEU LYS PHE ALA P1169 L1170 N1171 V1172 M1173 C1174 S1175 K1176 F1177 I1178
HIS CYS CYS PHE E1183 C1193 N1201 S1202 L1203 ASP ILE PRO MET LYS TYR PHE N1211 L1215 E1216 I1217 L1221 F1222 TYR VAL LEU ARG ASP THR ALA PRO GLU LEU PRO C1234 N1252 LYS GLU LEU THR ARG N1258 S1262 L1276 PRO ASN ALA ASN E1281 I1282 V1283 A1287
T1297 THR GLN VAL P1301 C1309 L1313 L1314 A1315 P1316 ILE PHE GLY LYS PRO LEU ARG ALA LEU PRO F1327 M1341 GLY LEU GLU ASN SER PHE LEU GLU TYR ASN E1352 L1357 V1358 Q1359 E1360 A1361 L1362 A1363 L1364 V1365 ASP ALA GLU ASP GLU SER LEU VAL SER ALA HIS ARG ILE SER
GLU HIS LYS THR S1384 I1402 THR LYS PRO GLU PHE ALA ALA ALA GLN GLN R1413 V1422 K1426 S1427 L1428 C1429 GLY ARG SER ILE E1434 I1435 A1439 H1440 G1441 K1444 A1445 V1446 I1447 ASP LEU LYS MET LYS LEU P1454 L1458 L1462 M1465 L1469 S1470 D1471 H1472 LYS LYS
L1475 A1482 S1492 Y1493 F1494 LYS V1496 G1497 L1515 Q1516 Q1517 LEU GLY SER GLN ASP LEU GLU ASN ASN S1527 T1528 V1529 Q1530 I1531 I1532 V1533 L1536 D1537 L1541 LEU PRO PRO THR A1546 F1549 M1550 A1556 Y1559 L1560 E1561 L1564 HIS ARG CYS GLN TYR SER P1571 F1572 R1573
E1574 P1575 L1576 A1577 K1578 F1579 L1580 D1581 ARG PHE P1584 N1593 S1596 LYS ARG E1599 I1600 T1601 T1602 R1603 Y1606 F1607 L1610 D1611 SER CYS S1614 A1618 L1621 E1622 R1626 L1630 L1631 H1632 GLN GLU GLY SER ALA GLU GLU LYS C1641 V1642 R1643 L1654 ALA ALA SER ASP LYS
GLU TRP ILE LYS ASP LYS E1666 E1667 L1668 L1669 G1670 E1671 L1672 L1673 D1674 A1675 G1676 SER V1678 C1679 L1680 R1684 V1688 VAL SER PRO LEU TYR F1694 S1713 GLN PRO L1716 L1733 PRO PHE VAL LEU GLU F1739 E1746 VAL VAL K1749 V1756 L1760 D1761 R1765 Q1769 V1770 SER
SER L1773 D1774 A1775 R1776 L1779 Y1780 I1783 F1784 E1791 H1795 G1796 ASP LEU SER ARG LEU SER GLN THR GLU ASN ASN GLU L1809 S1815 D1819 K1822 ALA THR GLY PRO LEU VAL ASP ASP TYR THR SER THR LEU GLU ASP R1838 E1842 L1846 L1849 LEU ILE K1852
K1867 N1871 Y1872 ILE LYS L1875 A1889 Y1890 S1893 R1894 PHE ASP T1897 F1906 V1907 A1908 L1909 L1910 R1911 C1912 H1913 GLN ILE D1916 Y1919 Q1923 A1924 L1925 E1926 L1927 L1928 A1929 S1933 GLU ARG THR ASN SER GLU LEU ASP TRP LEU LYS TRP PRO ARG ARG VAL LEU SER GLU





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	105916	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; Full CTF correction in Relion	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	60	Depositor
Minimum defocus (nm)	1400	Depositor
Maximum defocus (nm)	3400	Depositor
Magnification	127272	Depositor
Image detector	FEI FALCON II (4k x 4k)	Depositor
Maximum map value	0.192	Depositor
Minimum map value	-0.114	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.0429	Depositor
Map size (Å)	352.0, 352.0, 352.0	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.1, 1.1, 1.1	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	B	0.23	0/13920	0.40	3/19262 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	782	PRO	N-CA-CB	5.66	110.09	103.30
1	B	2717	ARG	O-C-N	5.64	131.73	122.70
1	B	3422	PHE	C-N-CA	5.40	135.21	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	3421	LEU	Peptide
1	B	3423	HIS	Peptide

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	B	14023	0	6129	270	0
All	All	14023	0	6129	270	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 13.

The worst 5 of 270 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:B:3651:HIS:O	1:B:3659:ILE:HA	1.02	1.20
1:B:3651:HIS:O	1:B:3659:ILE:CA	1.95	1.12
1:B:3420:SER:O	1:B:3423:HIS:CA	2.00	1.09
1:B:3420:SER:O	1:B:3423:HIS:HA	1.50	1.09
1:B:3420:SER:O	1:B:3423:HIS:N	1.95	0.99

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	B	2619/3825 (68%)	2264 (86%)	240 (9%)	115 (4%)	2 22

5 of 115 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	137	HIS
1	B	781	GLU
1	B	782	PRO
1	B	939	PRO
1	B	983	ASP

5.3.2 Protein sidechains

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

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 monosaccharides 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-3790. 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: 160

Y Index: 160

Z Index: 160

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: 161

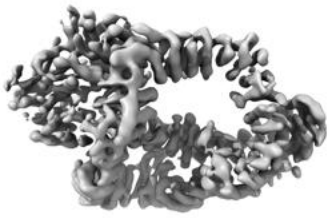
Y Index: 138

Z Index: 148

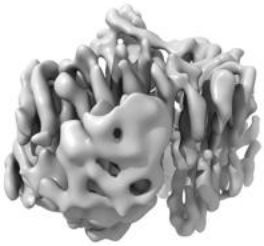
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

6.4.1 Primary map



X



Y

Z

The images above show the 3D surface view of the map at the recommended contour level 0.0429. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

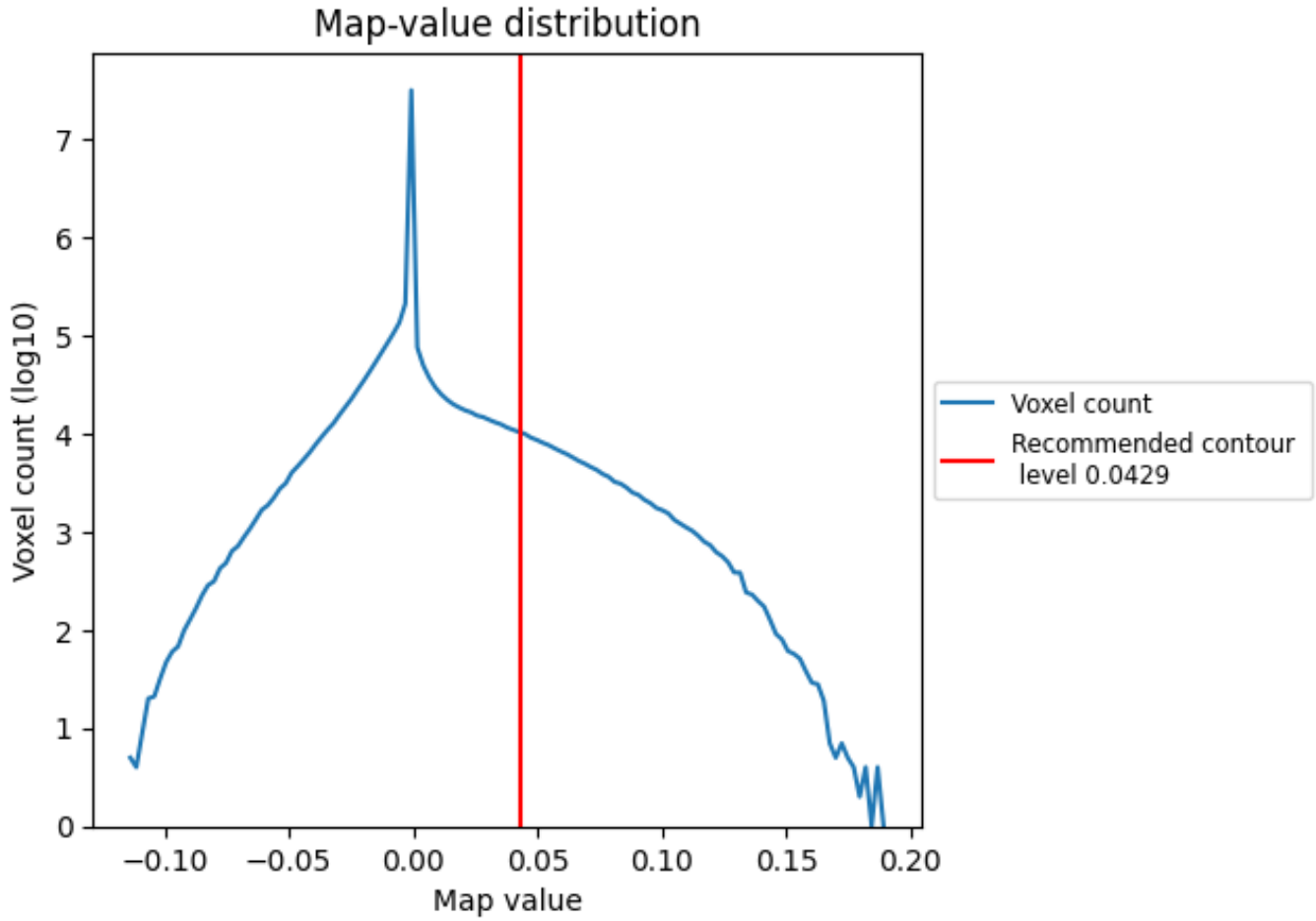
6.5 Mask visualisation (i)

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

7 Map analysis

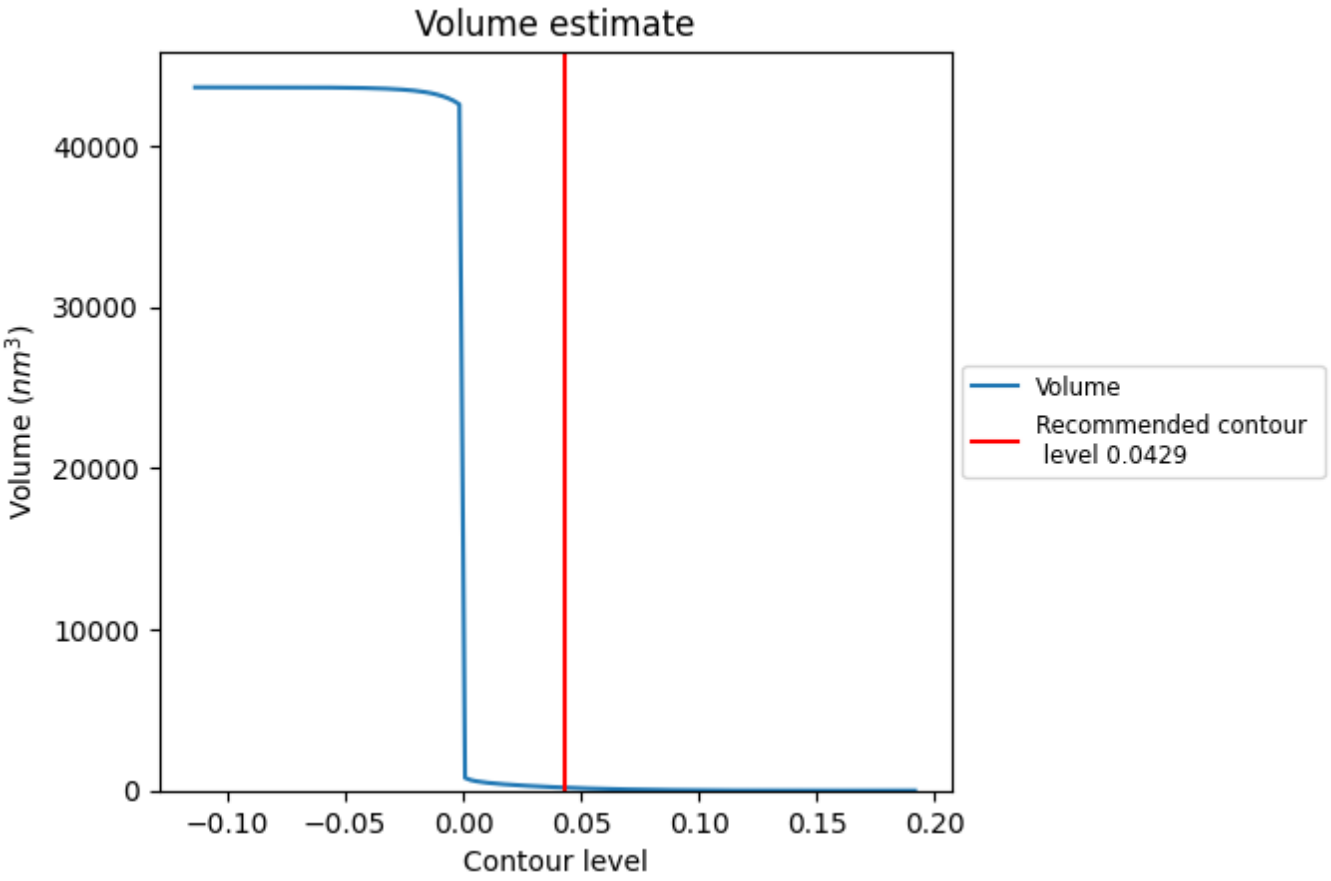
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

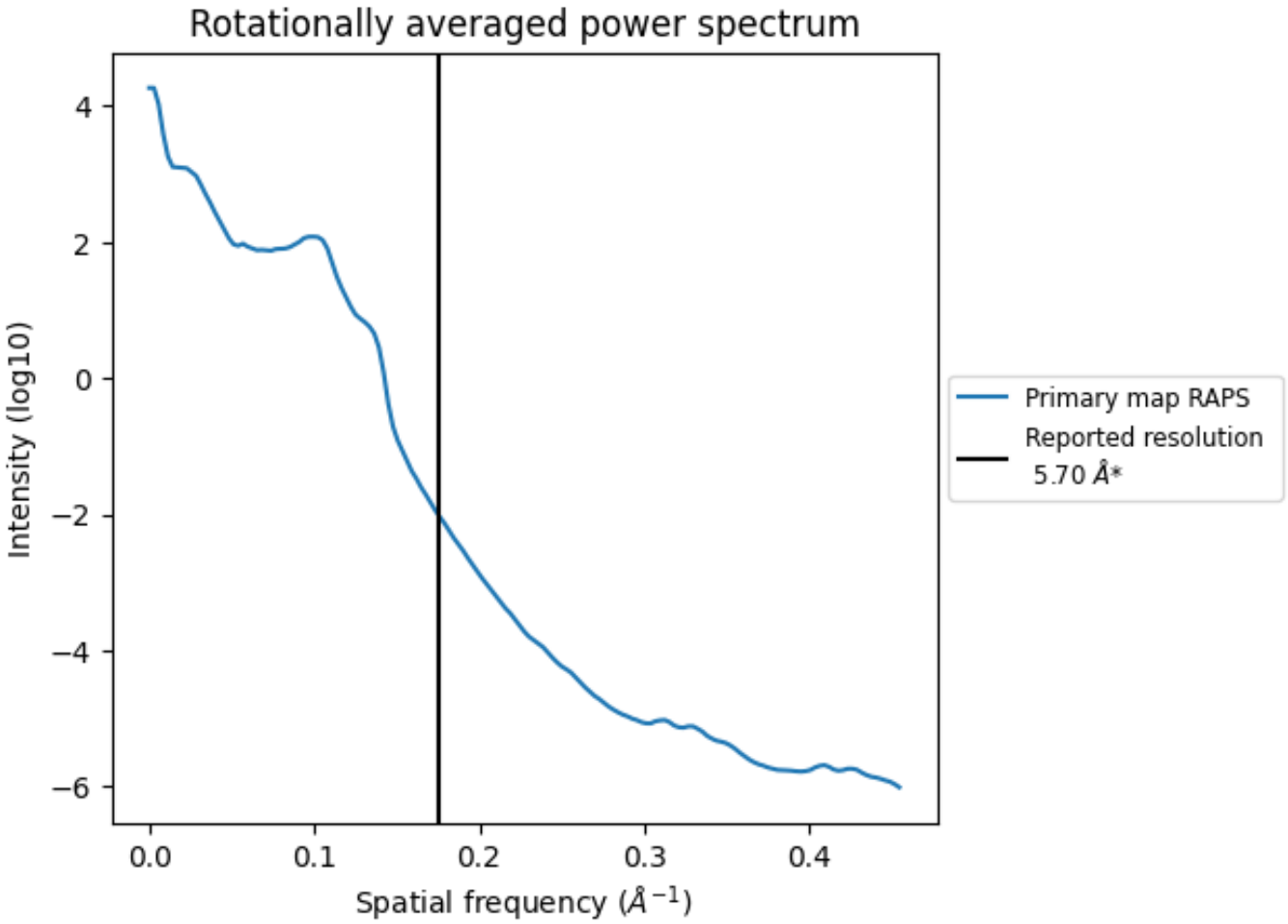
7.2 Volume estimate (i)



The volume at the recommended contour level is 184 nm^3; this corresponds to an approximate mass of 166 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)

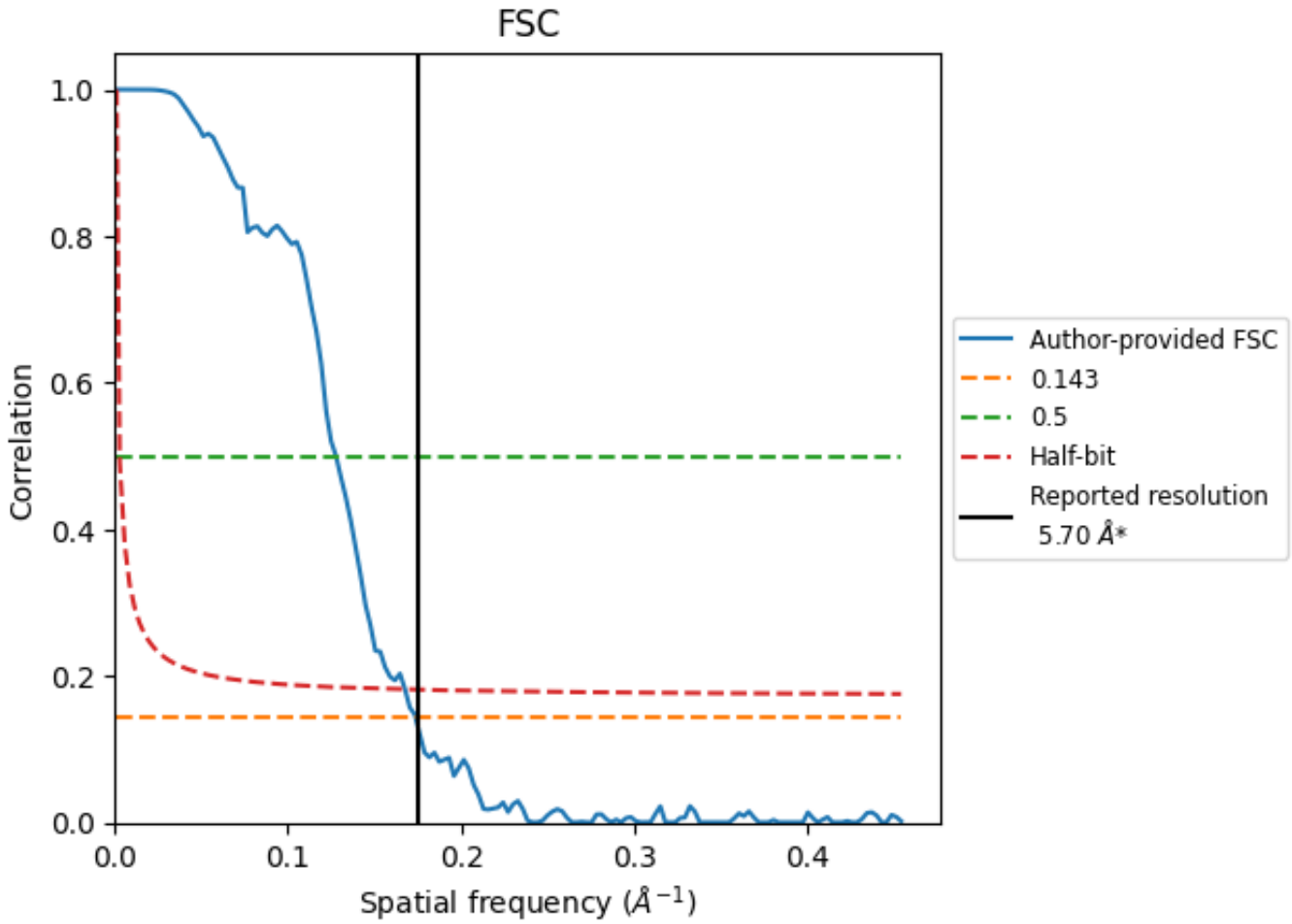


*Reported resolution corresponds to spatial frequency of 0.175 Å^{-1}

8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.175 Å^{-1}

8.2 Resolution estimates (i)

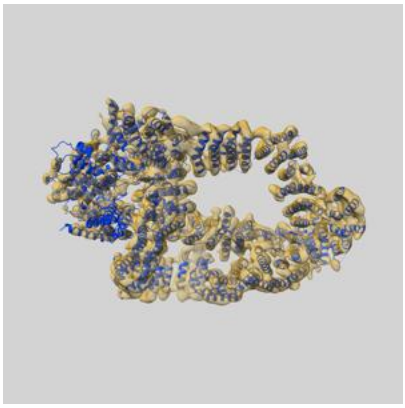
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	5.70	-	-
Author-provided FSC curve	5.75	7.82	5.96
Unmasked-calculated*	-	-	-

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

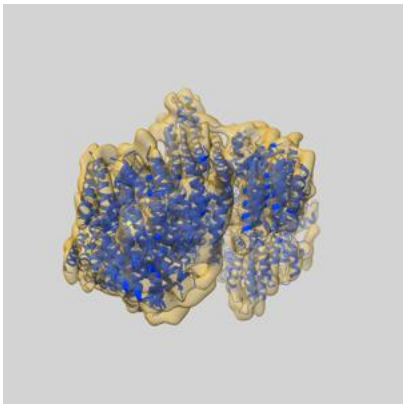
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-3790 and PDB model 5OEJ. Per-residue inclusion information can be found in section 3 on page 4.

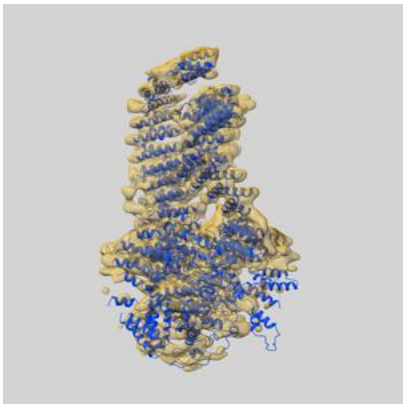
9.1 Map-model overlay



X



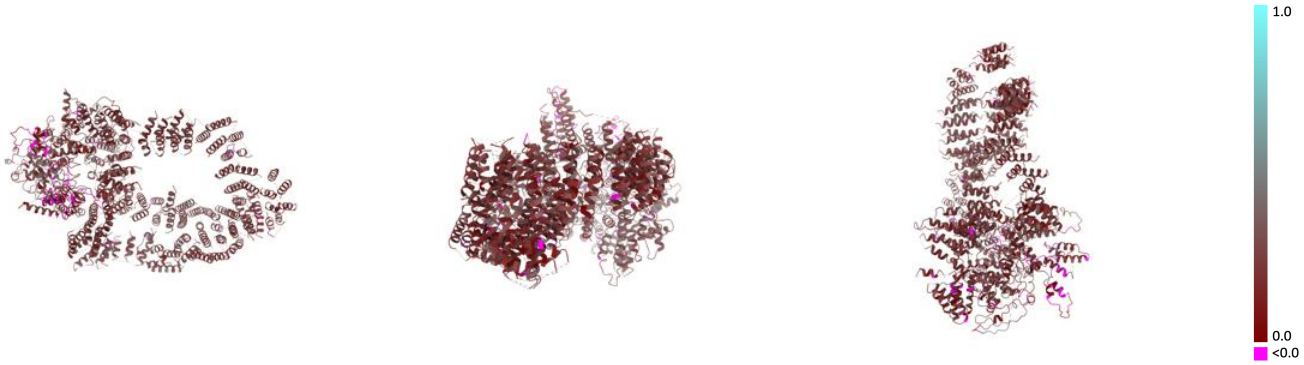
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0429 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.

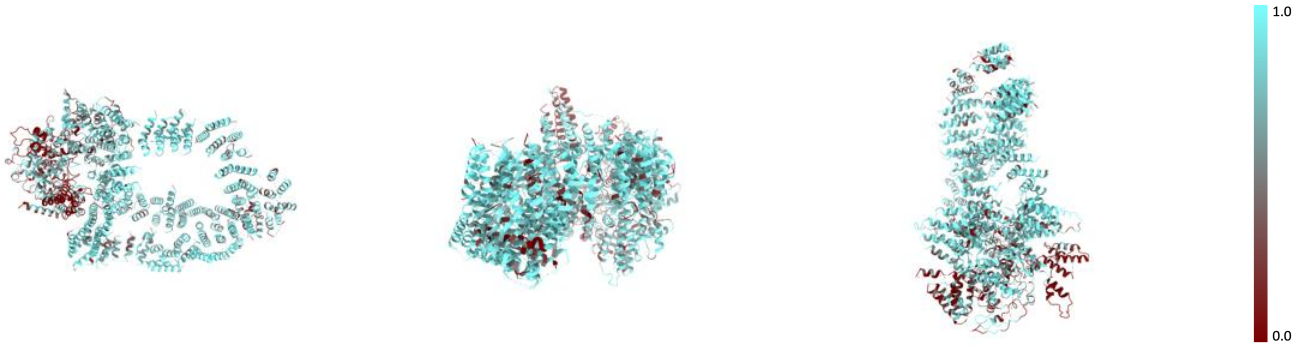
9.2 Q-score mapped to coordinate model



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.

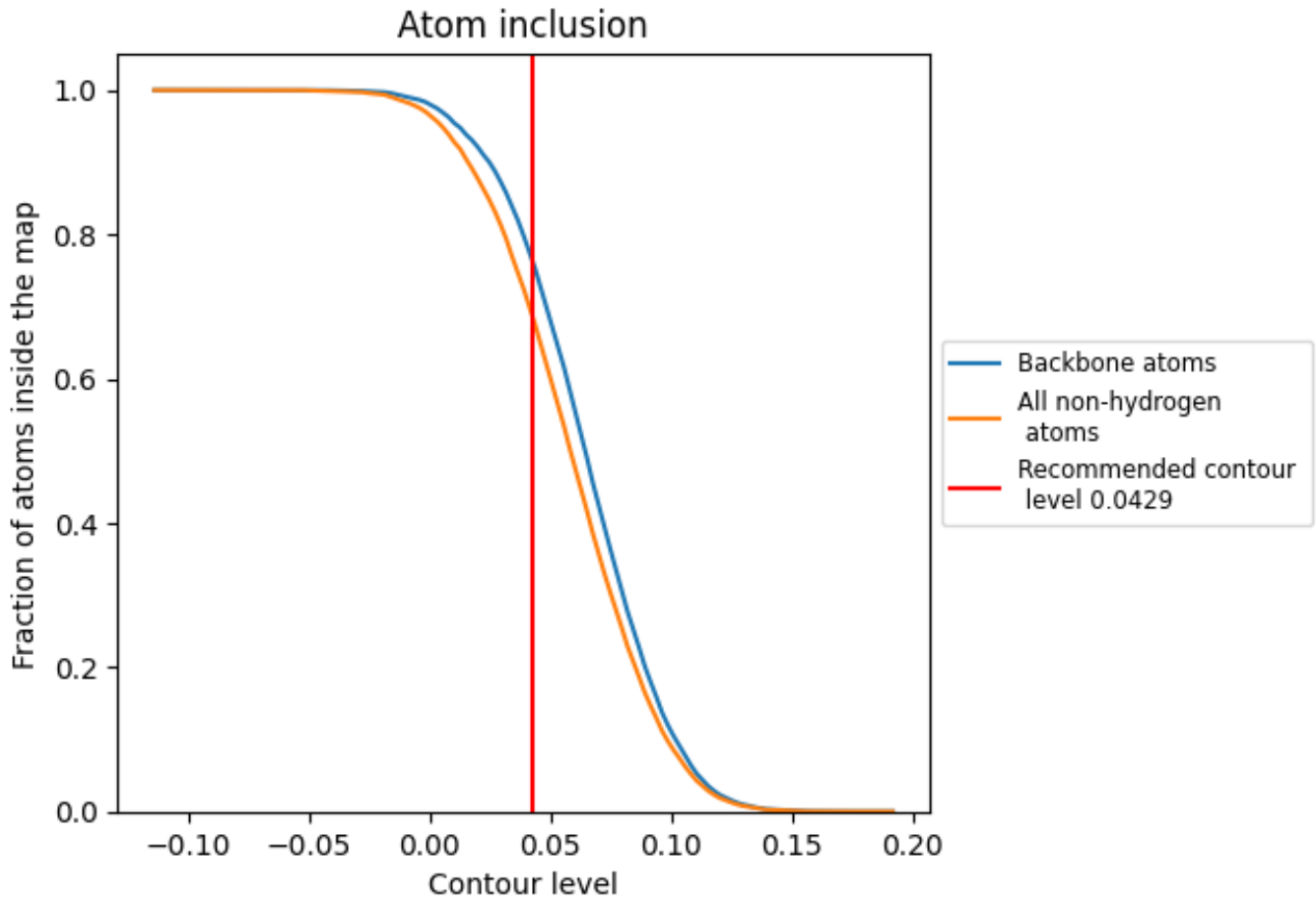
9.3 Atom inclusion mapped to coordinate model



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 (0.0429).

9.4 Atom inclusion (i)



At the recommended contour level, 76% of all backbone atoms, 68% 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 (0.0429) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6830	0.2190
B	0.6830	0.2190