wwPDB EM Validation Summary Report (i)

Dec 12, 2022 – 07:02 am GMT

PDB ID	:	5A5B
EMDB ID	:	EMD-3034
Title	:	Structure of the 26S proteasome-Ubp6 complex
Authors	:	Aufderheide, A.; Beck, F.; Stengel, F.; Hartwig, M.; Schweitzer, A.; Pfeifer, G.; Goldberg, A.L.; Sakata, E.; Baumeister, W.; Foerster, F.
Deposited on	:	2015-06-17
Resolution	:	9.50 Å(reported)
Based on initial models	:	1VJV, 4C43

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:

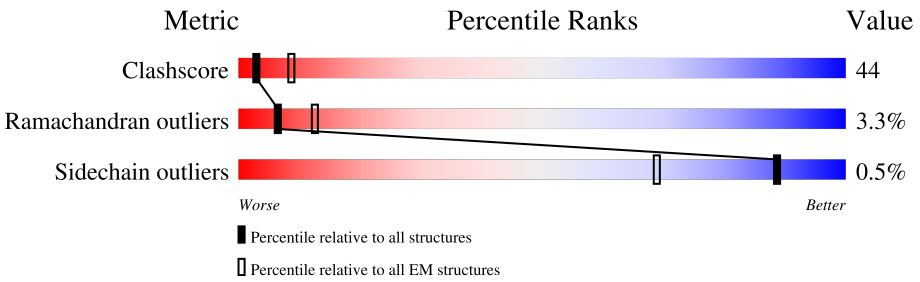
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 9.50 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	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	1	215	60%; 47%, 44%, • 5%
2	2	261	61%; 55%, 26%, 5%, 15%
3	3	205	71%; 65%, 33%, •
4	4	198	59%; 63%, 35%, •
5	5	287	40%; 40%, 31%, •, 26%
6	6	241	49%; 57%, 33%, •, 8%
7	7	266	48%; 48%, 37%, •, 12%
8	8	416	32%; 75%, 13%, 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	9	76	46% 62% 36%
10	A	252	27% 45% 47%
11	B	250	36% 57% 39%
12	C	258	33% 53% 41% 5%
13	D	254	22% 65% 27% 5%
14	E	260	27% 57% 36% 7%
15	F	234	22% 50% 44% 5%
16	G	288	23% 44% 37% 15%
17	H	467	34% 43% 29% 23%
18	I	437	33% 43% 37% 17%
19	J	405	29% 49% 40% 8%
20	K	428	35% 46% 40% 11%
21	L	437	27% 40% 40% 17%
22	M	434	38% 48% 33% 15%
23	N	945	24% 53% 33% 10%
24	O	393	20% 37% 54% 6%
25	P	445	12% 48% 40% 5% 7%
26	Q	434	18% 53% 43%
27	R	429	20% 46% 43% 5% 7%
28	S	523	25% 39% 26% 33%
29	T	274	27% 49% 42% 8%
30	U	338	31% 37% 31% 5% 25%
31	V	306	24% 38% 38% 19%
32	W	268	25% 29% 35% 7% 26%
33	X	156	37% 35% 40% 6% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
34	Y	89	11% 10% 79%
35	Z	993	33% 41% 36% 5% 18%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
9	GLZ	9	76	-	-	X	-

2 Entry composition (i)

There are 35 unique types of molecules in this entry. The entry contains 83748 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 PROTEASOME COMPONENT PRE3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	1	205	1576	996	261	312	7	0	0

- Molecule 2 is a protein called PROTEASOME COMPONENT PUP1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	2	223	1692	1067	294	324	7	0	0

- Molecule 3 is a protein called PROTEASOME COMPONENT PUP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	3	204	1581	1010	258	305	8	0	0

- Molecule 4 is a protein called PROTEASOME COMPONENT C11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	4	198	1585	1005	269	305	6	0	0

- Molecule 5 is a protein called PROTEASOME COMPONENT PRE2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	5	212	1646	1045	282	312	7	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
5	33	ARG	LYS	conflict	UNP P30656

- Molecule 6 is a protein called PROTEASOME COMPONENT C5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	6	222	1757	1115	303	335	4	0	0

- Molecule 7 is a protein called PROTEASOME COMPONENT PRE4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	7	233	1824	1154	312	351	7	0	0

- Molecule 8 is a protein called UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 6.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	S	Se		
8	8	368	3008	1922	500	571	5	10	10	1

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
8	85	MSE	-	expression tag	UNP P43593
8	86	GLY	-	expression tag	UNP P43593
8	87	SER	-	expression tag	UNP P43593
8	88	ASP	-	expression tag	UNP P43593
8	89	LYS	-	expression tag	UNP P43593
8	90	ILE	-	expression tag	UNP P43593
8	91	HIS	-	expression tag	UNP P43593
8	92	HIS	-	expression tag	UNP P43593
8	93	HIS	-	expression tag	UNP P43593
8	94	HIS	-	expression tag	UNP P43593
8	95	HIS	-	expression tag	UNP P43593
8	96	HIS	-	expression tag	UNP P43593
8	500	MSE	-	expression tag	UNP P43593
8	182	VAL	ILE	conflict	UNP P43593

- Molecule 9 is a protein called UBIQUITIN ALDEHYDE.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	9	76	601	378	105	117	1	0	0

- Molecule 10 is a protein called PROTEASOME COMPONENT C7-ALPHA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	A	243	1921	1221	322	370	8	0	0

- Molecule 11 is a protein called PROTEASOME COMPONENT Y7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	B	250	1915	1219	315	377	4	0	0

- Molecule 12 is a protein called PROTEASOME COMPONENT Y13.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	C	245	1913	1207	323	380	3	0	0

- Molecule 13 is a protein called PROTEASOME COMPONENT PRE6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	D	242	1899	1186	333	376	4	0	0

- Molecule 14 is a protein called PROTEASOME COMPONENT PUP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	E	243	1867	1165	315	380	7	0	0

- Molecule 15 is a protein called PROTEASOME COMPONENT PRE5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	F	233	1795	1129	312	350	4	0	0

- Molecule 16 is a protein called PROTEASOME COMPONENT C1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	G	245	1900	1207	331	358	4	0	0

- Molecule 17 is a protein called 26S PROTEASE REGULATORY SUBUNIT 7 HOMOLOG.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	H	359	2792	1755	499	523	15	0	0

- Molecule 18 is a protein called 26S PROTEASE REGULATORY SUBUNIT 4 HOMOLOG.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	I	362	2822	1773	471	563	15	0	0

- Molecule 19 is a protein called 26S PROTEASE REGULATORY SUBUNIT 8 HOMOLOG.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	J	373	2928	1837	527	547	17	0	0

- Molecule 20 is a protein called 26S PROTEASE REGULATORY SUBUNIT 6B HOMOLOG.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	K	381	3019	1898	530	581	10	0	0

- Molecule 21 is a protein called 26S PROTEASE SUBUNIT RPT4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	L	361	2853	1798	507	536	12	0	0

- Molecule 22 is a protein called 26S PROTEASE REGULATORY SUBUNIT 6A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	M	367	2866	1799	503	553	11	0	0

- Molecule 23 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	N	849	6562	4174	1099	1261	28	0	0

- Molecule 24 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	O	387	3182	2047	520	606	9	0	0

- Molecule 25 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	P	415	3401	2166	571	655	9	0	0

- Molecule 26 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	Q	431	3471	2205	574	676	16	0	0

- Molecule 27 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	R	400	3218	2051	527	630	10	0	0

- Molecule 28 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	S	353	2893	1857	482	541	13	0	0

- Molecule 29 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN12.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	T	272	2235	1432	355	441	7	0	0

- Molecule 30 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	U	255	2061	1312	352	391	6	0	0

- Molecule 31 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	V	247	1942	1225	328	376	13	0	0

- Molecule 32 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN10.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
32	W	197	1534	962	269	300	3	0	0

- Molecule 33 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN13.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
33	X	127	1032	664	169	195	4	0	0

- Molecule 34 is a protein called 26S PROTEASOME COMPLEX SUBUNIT SEM1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
34	Y	19	168	101	30	37	0	0

- Molecule 35 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN1.

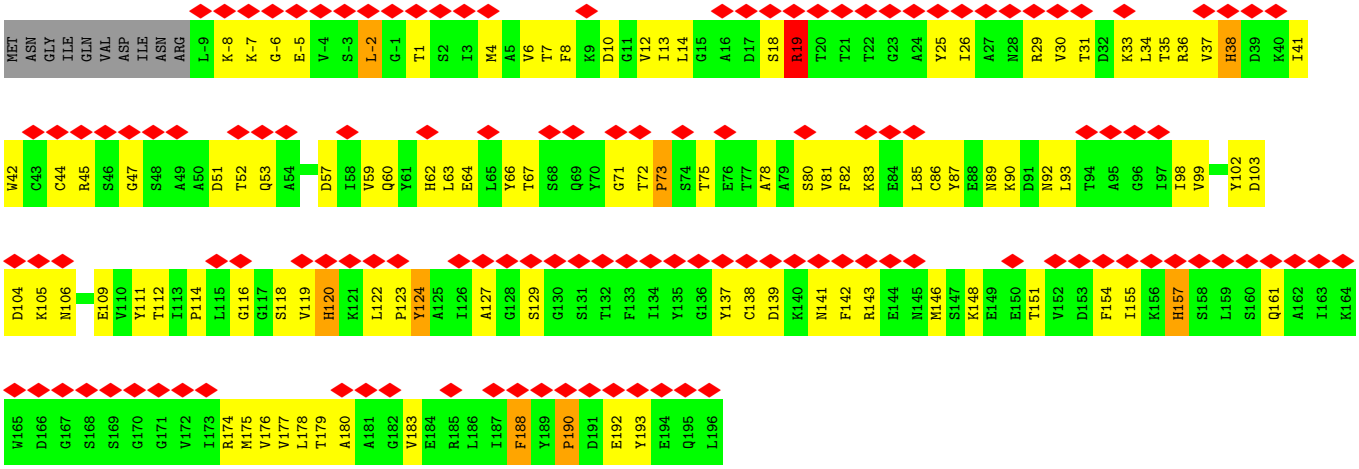
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
35	Z	813	6289	3995	1029	1236	29	0	0

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.

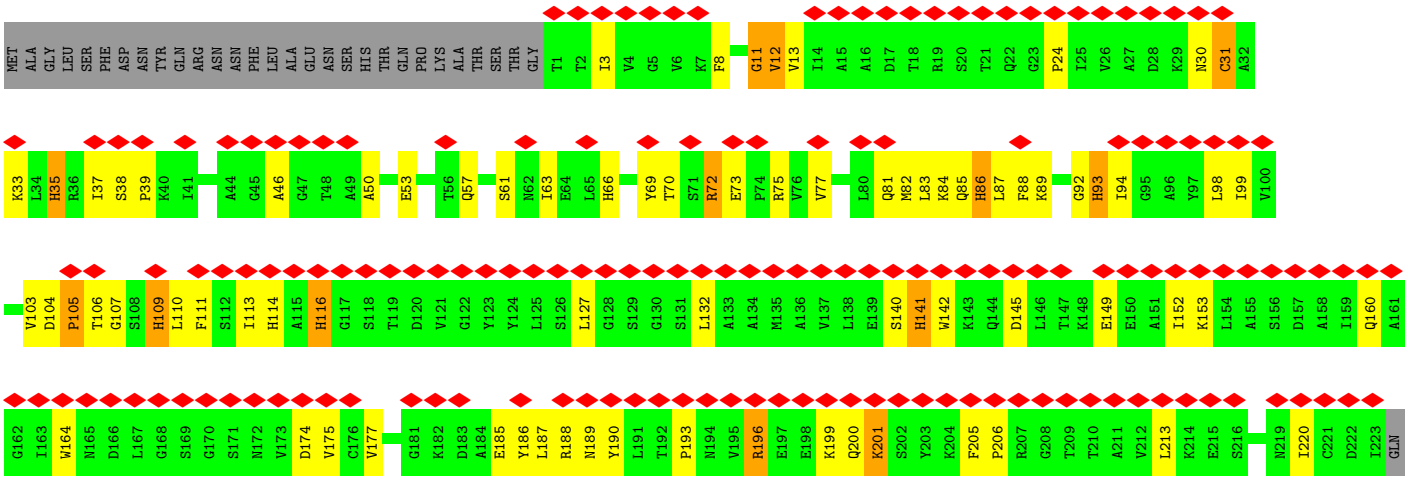
• Molecule 1: PROTEASOME COMPONENT PRE3

Chain 1: 60% 47% 44% • 5%



• Molecule 2: PROTEASOME COMPONENT PUP1

Chain 2: 61% 55% 26% 5% 15%



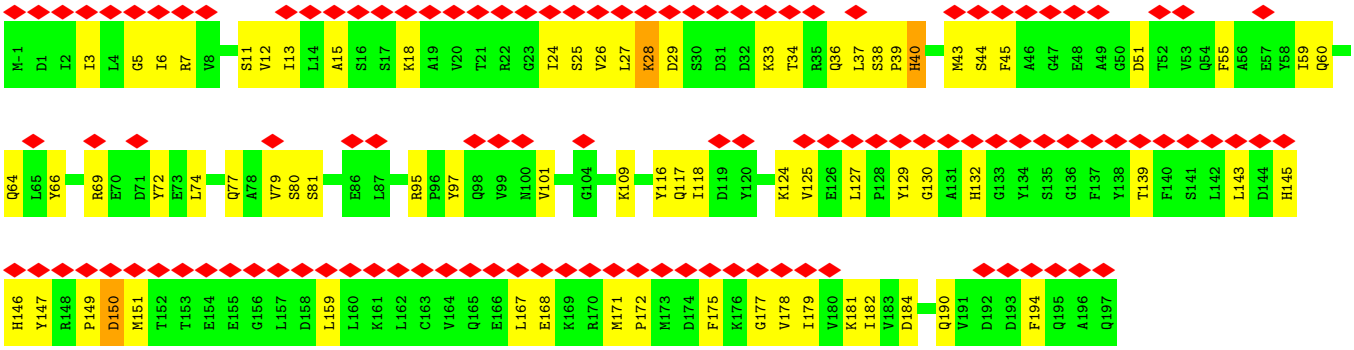
GLU GLU GLN VAL ASP ILE THR ALA

- Molecule 3: PROTEASOME COMPONENT PUP3

Chain 3:

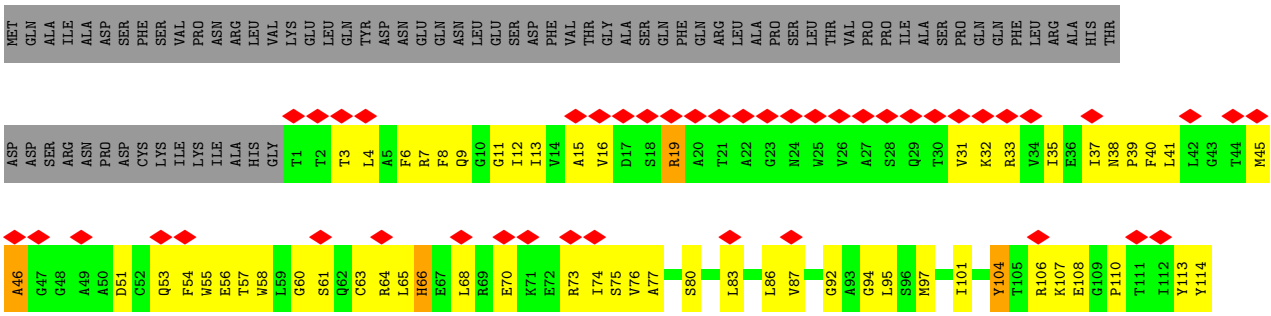
- Molecule 4: PROTEASOME COMPONENT C11

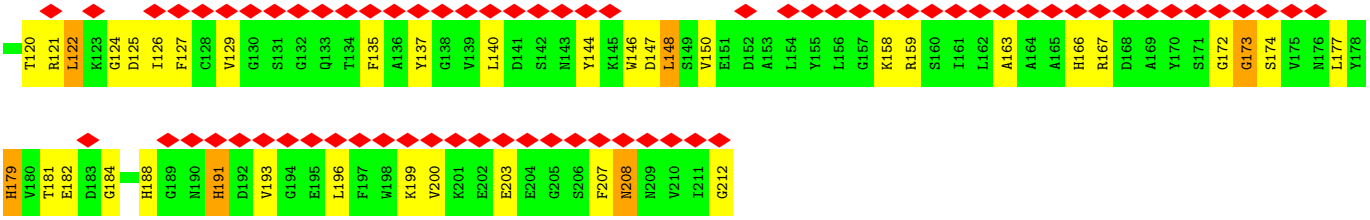
Chain 4:



- Molecule 5: PROTEASOME COMPONENT PRE2

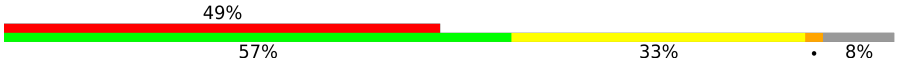
Chain 5:

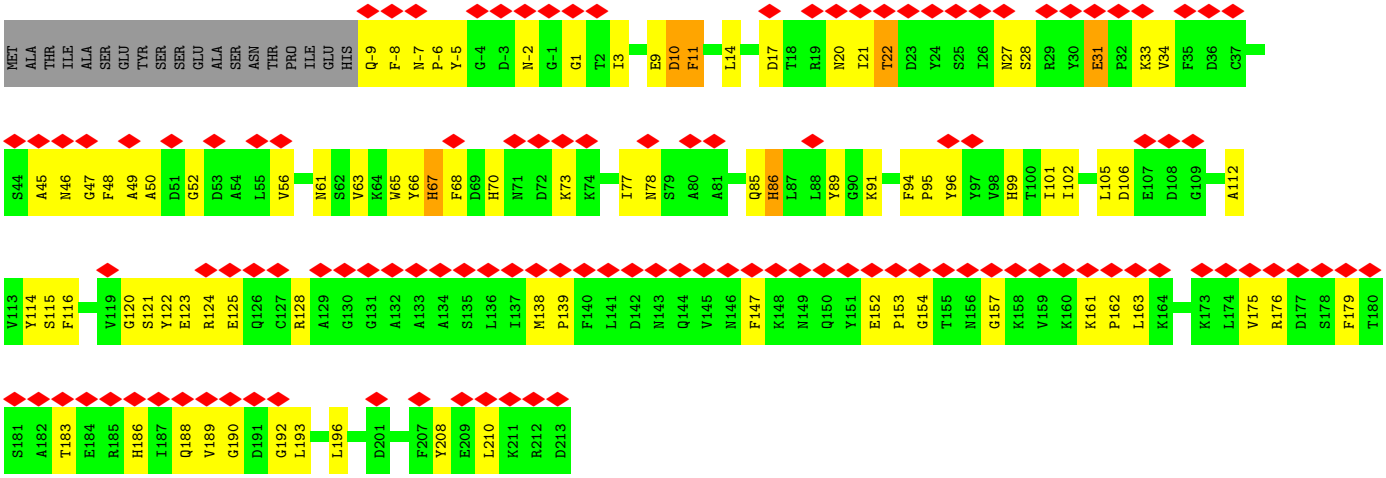




- Molecule 6: PROTEASOME COMPONENT C5

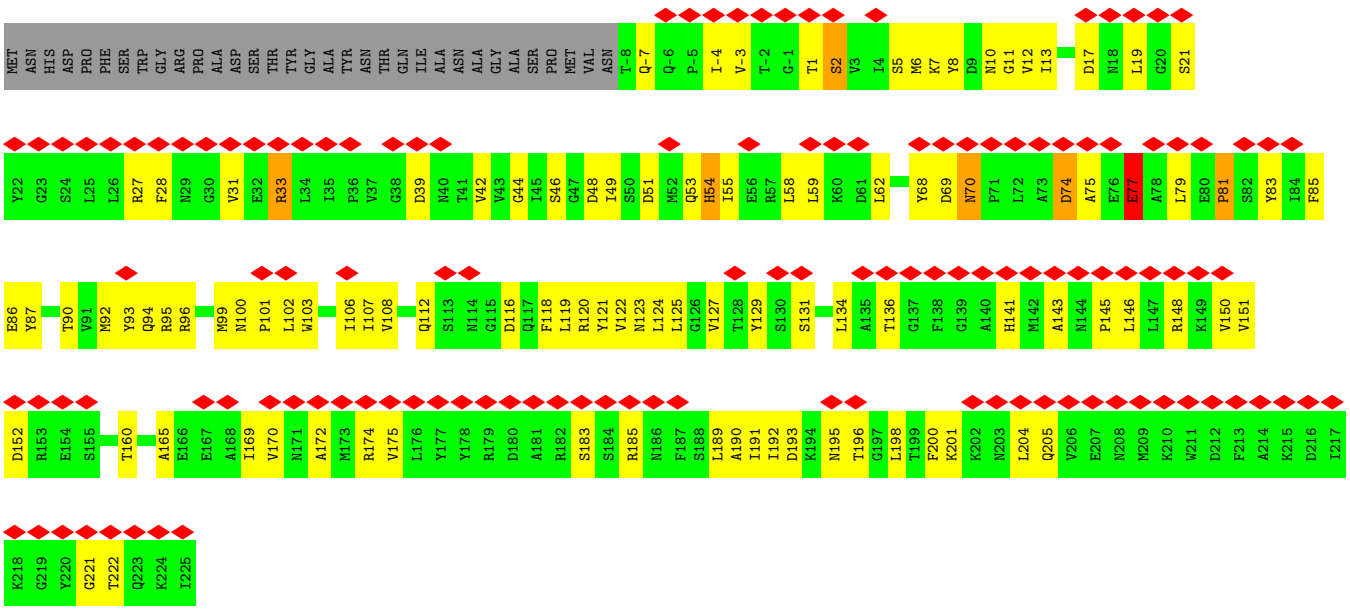
Chain 6:





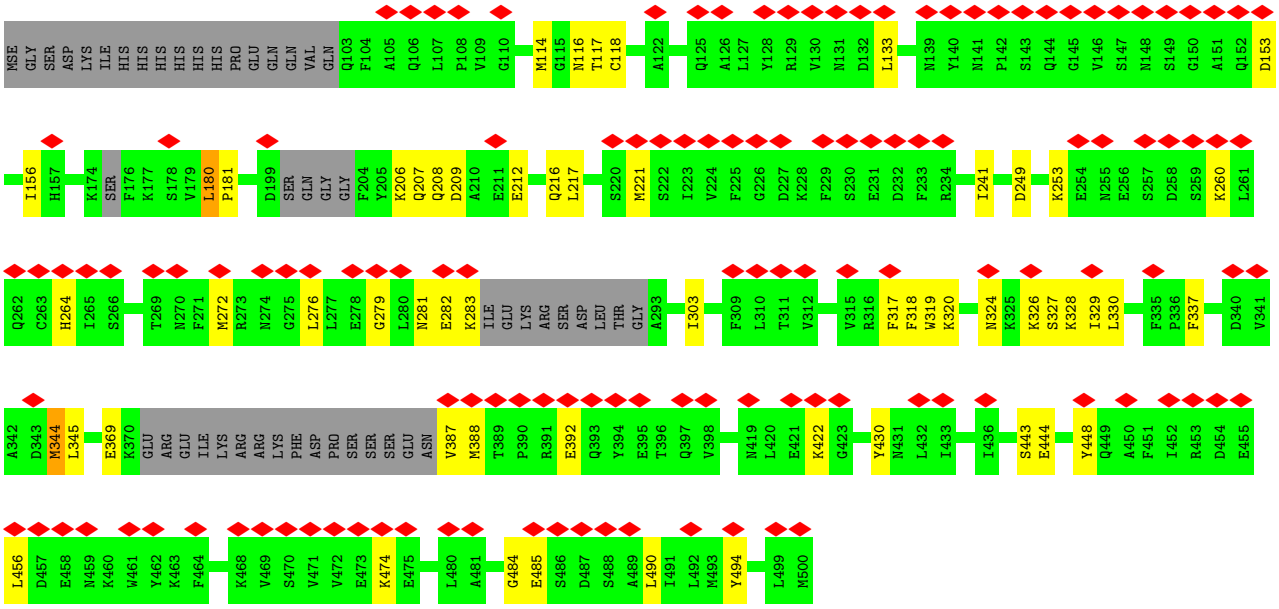
- Molecule 7: PROTEASOME COMPONENT PRE4

Chain 7:



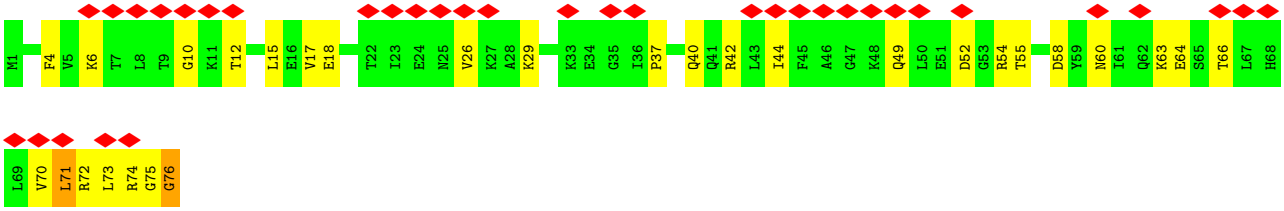
- Molecule 8: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 6

Chain 8:



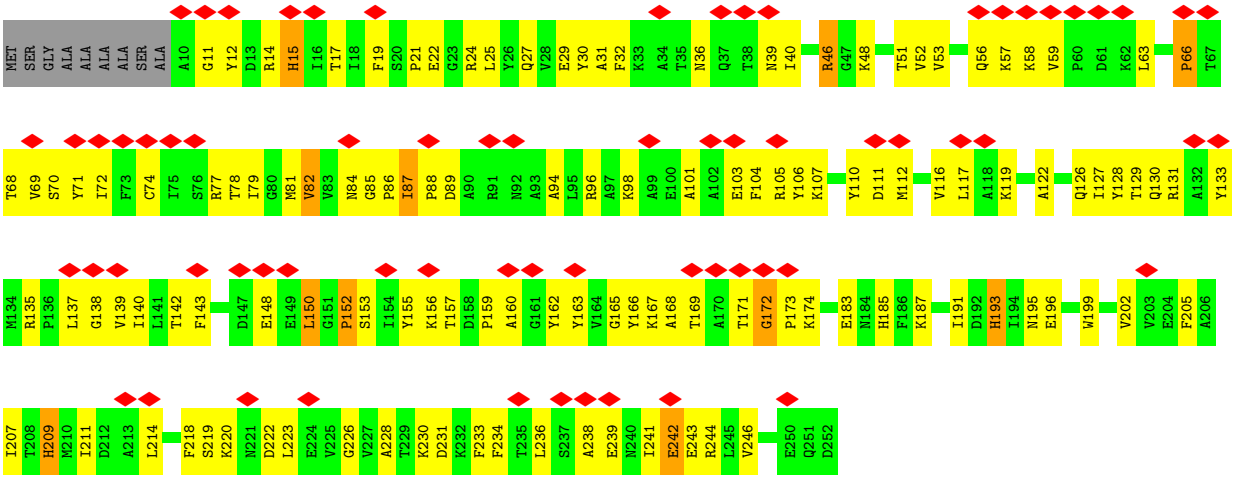
- Molecule 9: UBIQUITIN ALDEHYDE

Chain 9:



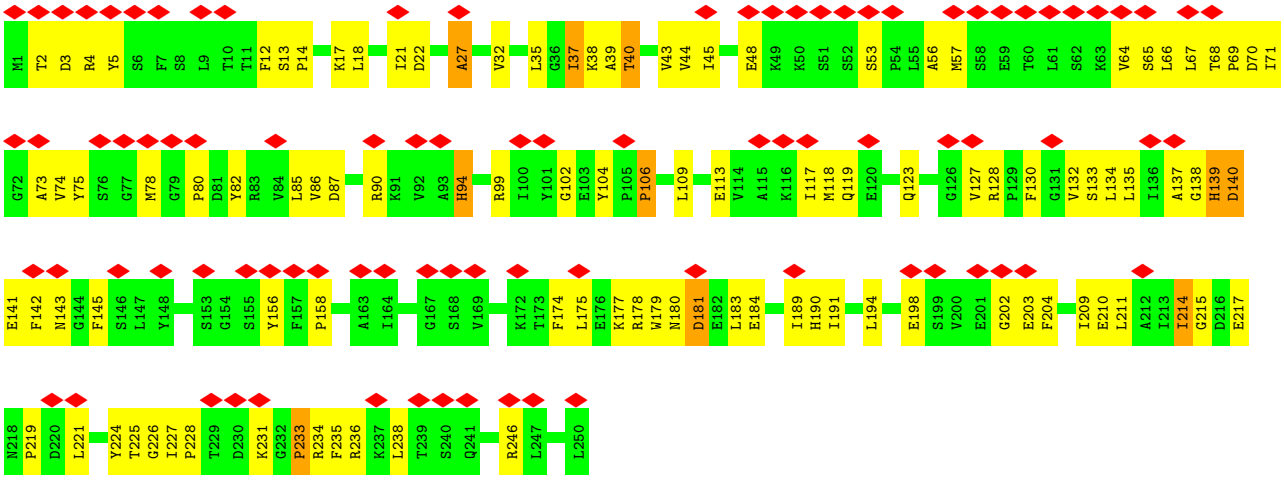
- Molecule 10: PROTEASOME COMPONENT C7-ALPHA

Chain A:



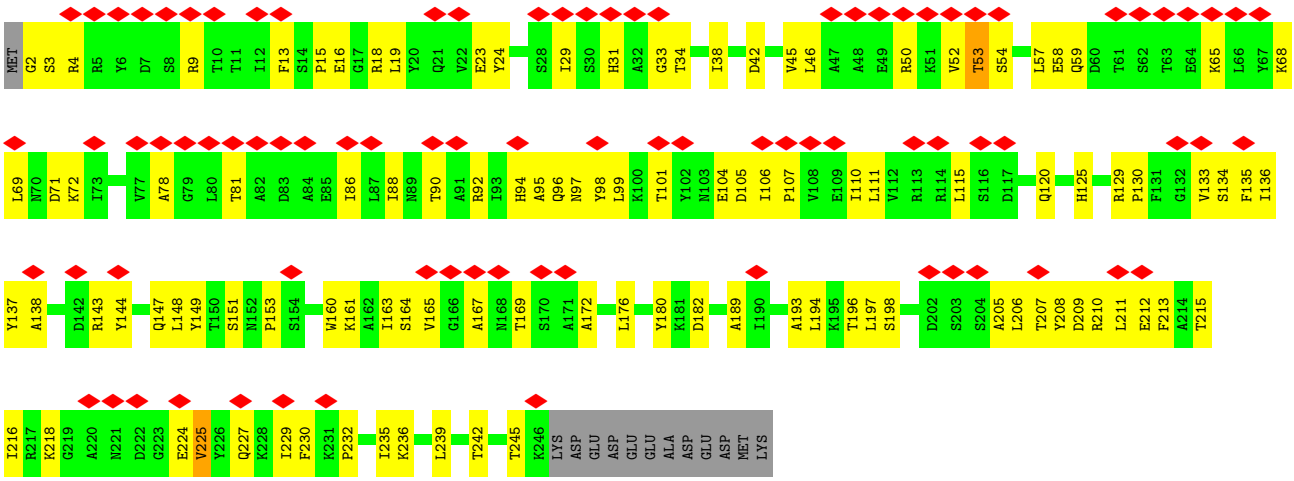
- Molecule 11: PROTEASOME COMPONENT Y7

Chain B: 36% 57% 39% •



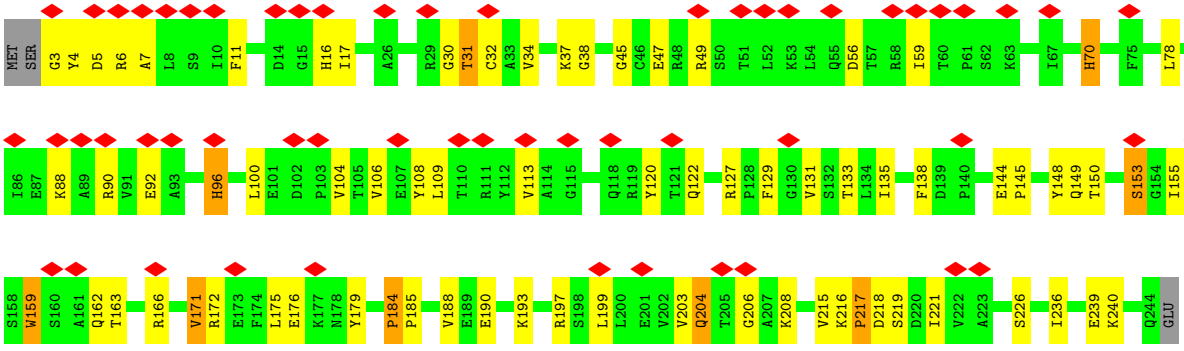
• Molecule 12: PROTEASOME COMPONENT Y13

Chain C: 33% 53% 41% • 5%



• Molecule 13: PROTEASOME COMPONENT PRE6

Chain D: 22% 65% 27% • 5%



GLN ASP LYS LYS LYS LYS SER ASN HIS

- Molecule 14: PROTEASOME COMPONENT PUP2

Chain E:

MET PHE LEU THR ARG SER GLU TYR D9 R10 G11 V12 F15 S16 P17 E18 G19 Y26 S27 L33 G34 S35 T36 A37 I38 T42 V46 V47 L48 G49 V50 E51 K52 R53 A54 T55 S56 P57 L58 L59 I64 E65 K66 I67 V68 E69 I70 D71 R72 H73 A77 M78

L81 T82 A83 D84 A85 M88 I89 E90 H91 A92 R93 V97 T98 H99 N100 L101 Y102 Y103 D104 E105 D106 I107 N108 V109 E110 S111 L112 T113 Q114 C117 L121 R122 F123 G124 E125 G126 A127 S128 G129 E130 E131 R132 L133 M134 S135 R136 P137 F138 G139 L142 L143 I144 A145 G146 H147

D148 A149 D150 Y153 Q154 L155 F156 H157 A158 E159 P160 S161 G162 T163 F164 Y165 R166 Y167 K170 A171 I172 G173 S174 G175 S176 E177 G178 A179 Q180 A181 E182 L183 L184 N185 H188 A196 V200 L201 L204 M208 E209 E210 K211 L212 N215 Q218 L219 I222 T223

D226 K229 I230 Y231 K235 T236 L239 I240 K241 E242 L243 K244 A249 E250 S251 PRO GLU GLU ALA ASP VAL GLU MET SER

- Molecule 15: PROTEASOME COMPONENT PRE5

Chain F:

MET F2 R3 N4 N5 Y6 D7 G8 D9 T10 V11 T12 F13 S14 P15 T16 G17 R18 Q21 Y24 Q31 G32 S33 V34 T35 L38 R39 S40 N41 T42 H43 A44 V45 L46 L49 K50 R51 E55 Y59 Q60 K61 K62 I63 I64 K65 C66 H69 L72 S73 L74 A75

G76 L77 A78 P79 D80 A81 R82 V83 L84 S85 N86 Y87 L88 R89 Q90 Q91 C92 N93 F99 N100 R101 K102 L103 A104 V105 E106 R107 A108 G109 H110 L111 L112 C113 D114 K115 A116 Q117 T120 Q121 S122 Y123 G124 G125 R126 P127 Y128 G129 V130 G131 L132 L133 I134 I135 G136 Y137 D138 K139 S140

H143 L144 L145 E146 F147 Q148 P149 S150 V153 T154 E155 L156 T159 A163 R164 S165 Q166 K169 T170 Y171 L172 E173 R174 F179 I180 K181 I182 P186 D187 E188 L189 I190 K191 A192 G193 V194 E195 A196 I197 S198 L201 R202 D203 E204 S205 L206 N210 L211 S212 I215

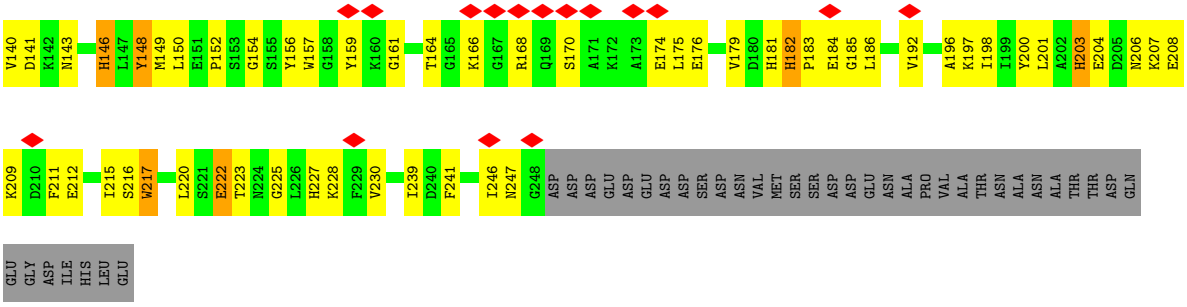
V216 G217 K218 D219 T220 P221 F222 G227 V230 Y233 I234

- Molecule 16: PROTEASOME COMPONENT C1

Chain G:

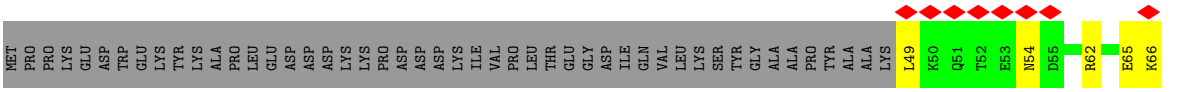
MET THR SER ILE G4 T5 G6 Y7 D8 L9 S10 N11 S12 V13 F14 D17 E24 Y25 K28 N32 G33 T34 T35 S36 I37 G38 I39 N42 V45 V46 F47 E50 I53 T54 S55 K56 L57 L58 V59 P60 Q61 K62 N63 V64 K65 I66 Q67 V68 V69 H72 I73

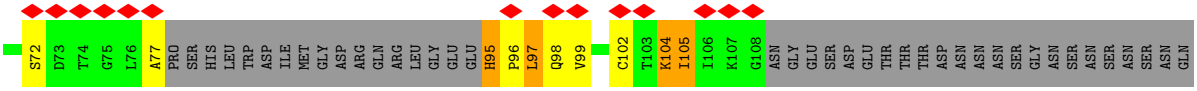
G74 C75 V76 Y77 S78 G79 L80 I81 P82 R85 H86 L87 V88 N89 R90 G91 R92 E93 E94 A95 A96 S97 F98 K103 T104 P105 I106 P107 I108 P109 A110 F111 A112 D113 R114 L115 G116 Q117 Y118 V119 Q120 A121 H122 T123 L124 Y125 N126 S127 V128 R129 P130 F131 S134 F137 G138 G139

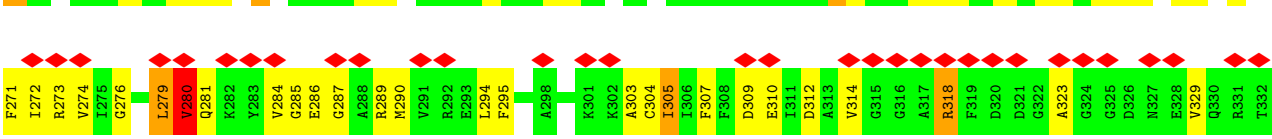


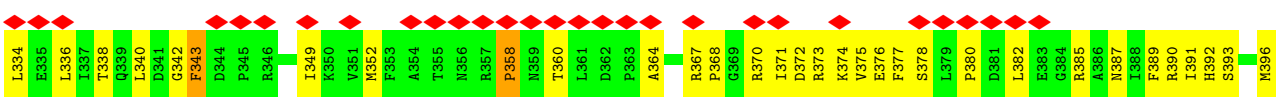
- Molecule 17: 26S PROTEASE REGULATORY SUBUNIT 7 HOMOLOG

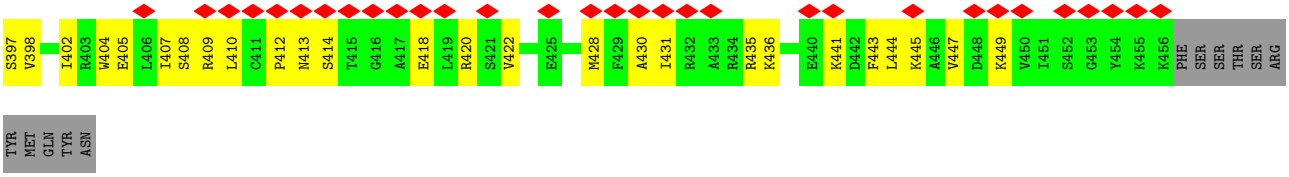
Chain H:





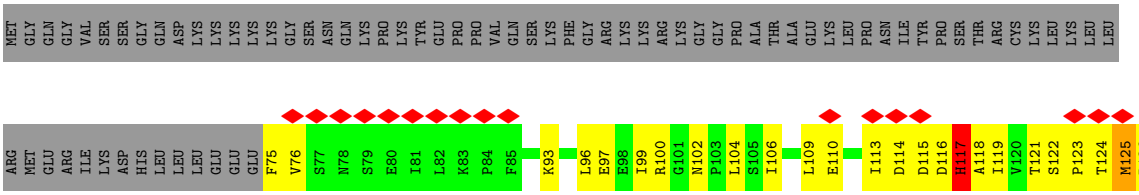


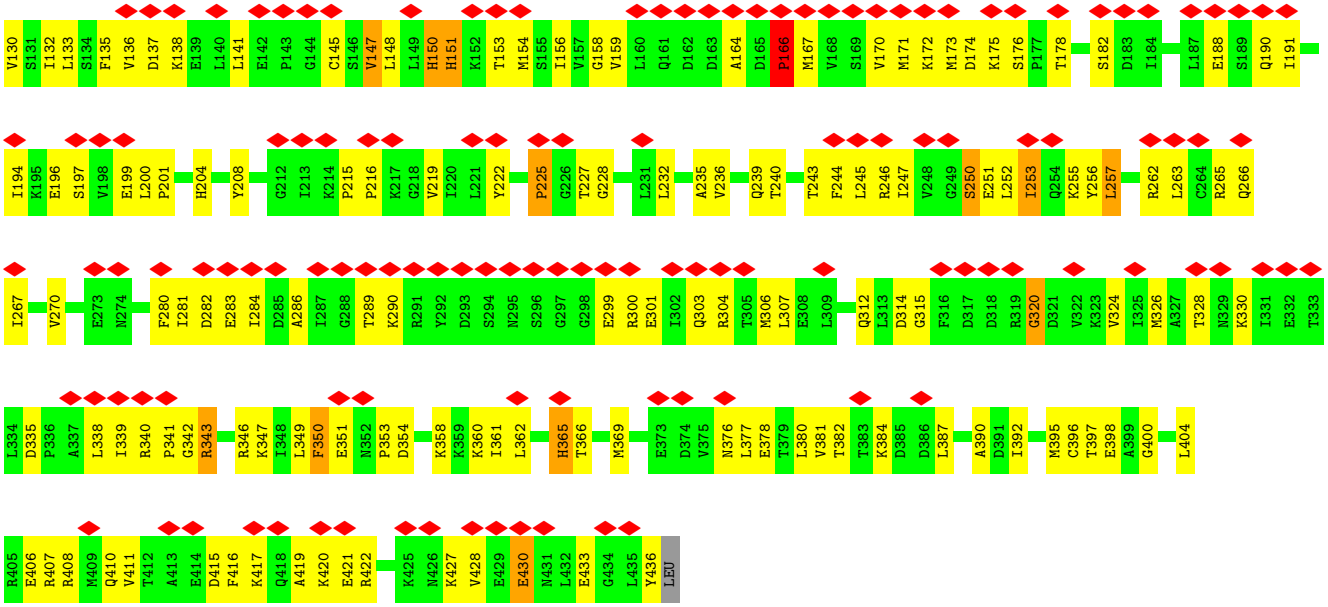




- Molecule 18: 26S PROTEASE REGULATORY SUBUNIT 4 HOMOLOG

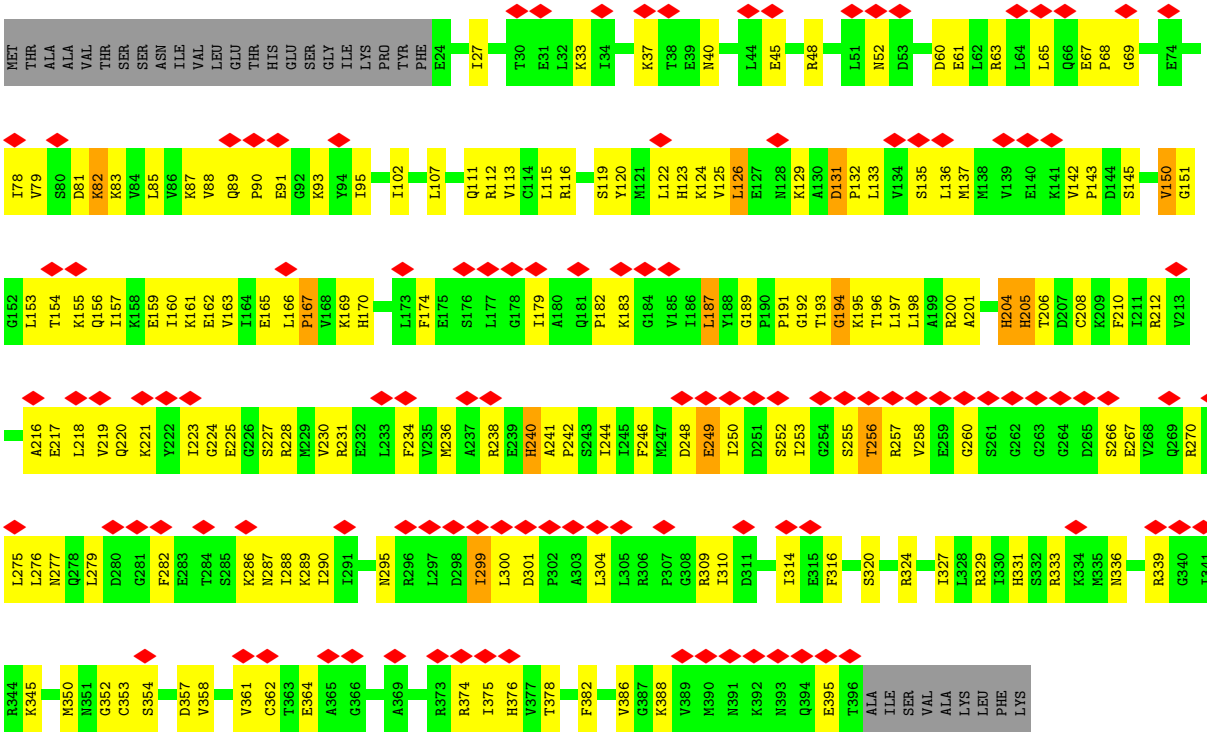
Chain I:





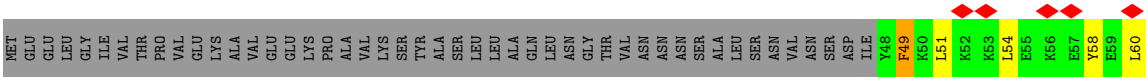
- Molecule 19: 26S PROTEASE REGULATORY SUBUNIT 8 HOMOLOG

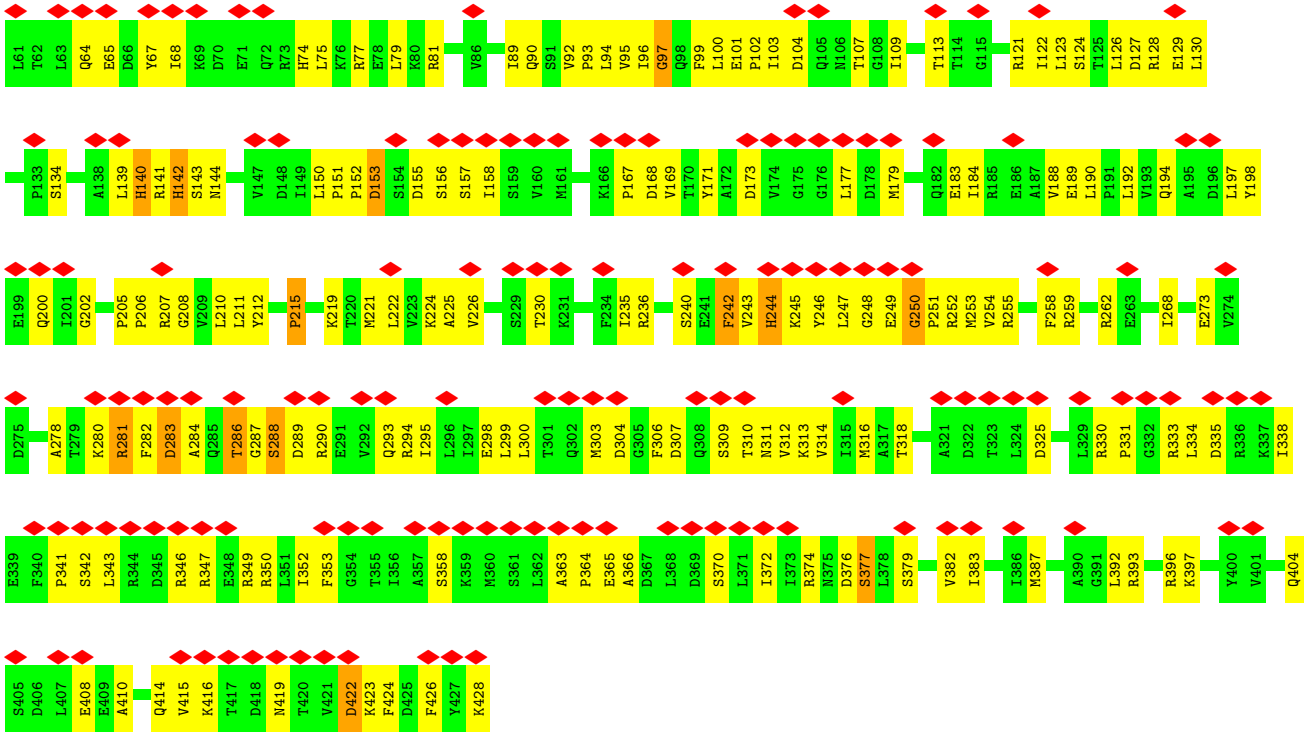
Chain J:



- Molecule 20: 26S PROTEASE REGULATORY SUBUNIT 6B HOMOLOG

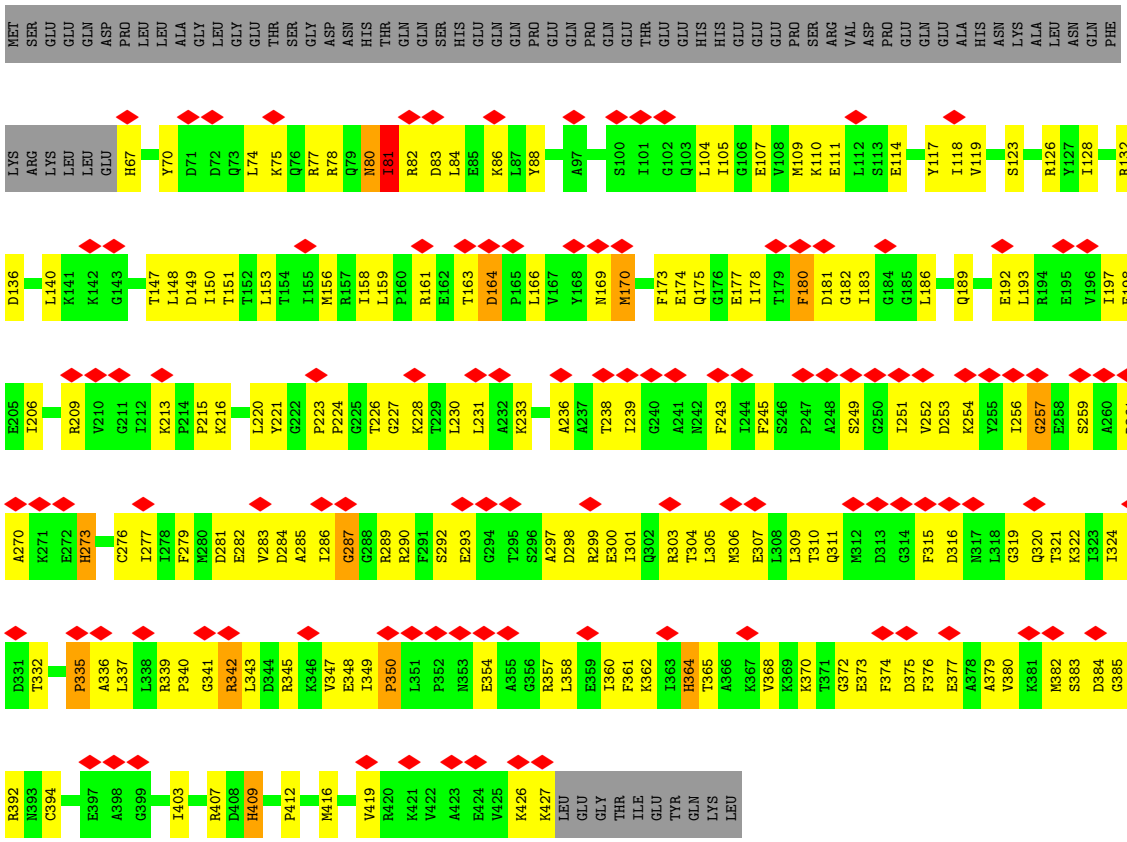
Chain K:





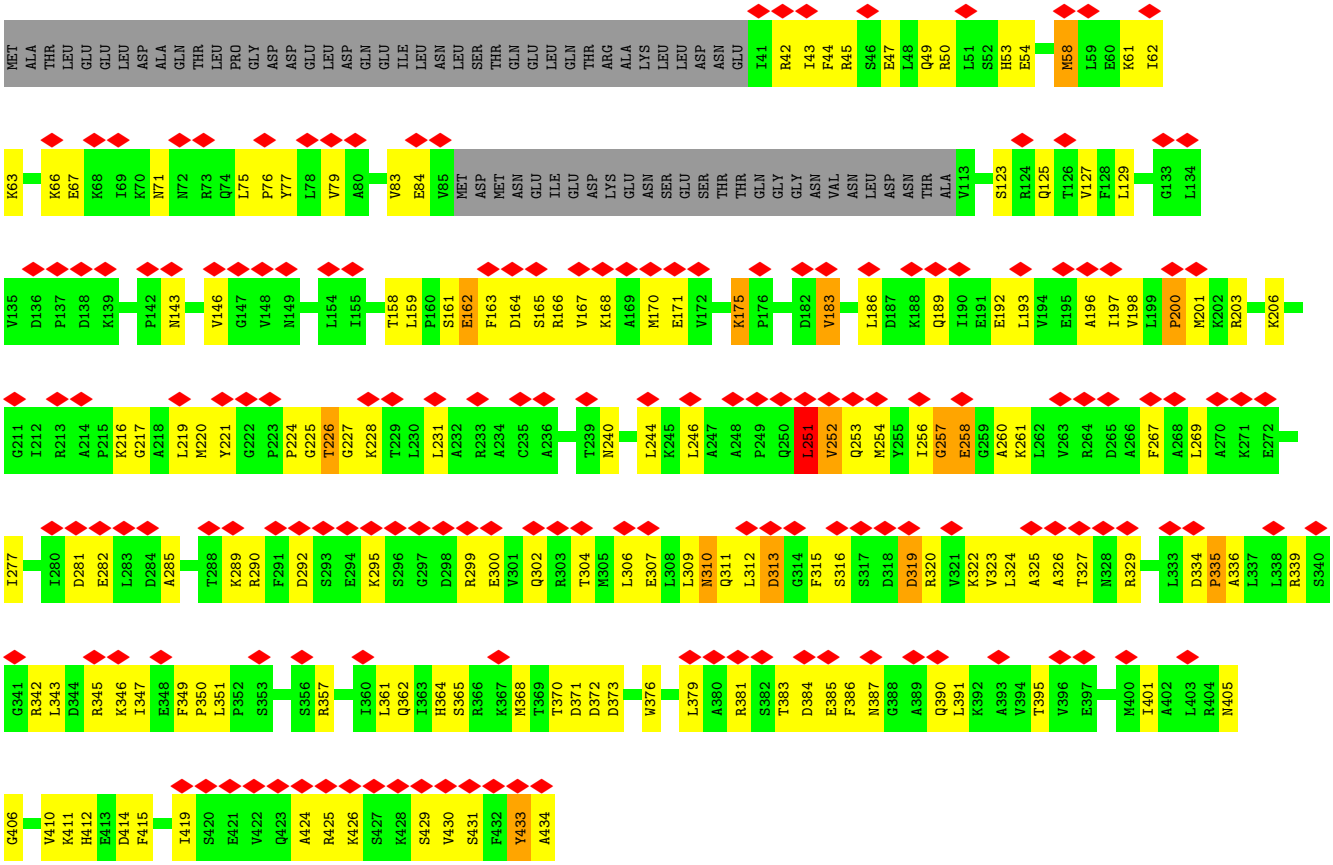
- Molecule 21: 26S PROTEASE SUBUNIT RPT4

Chain L:



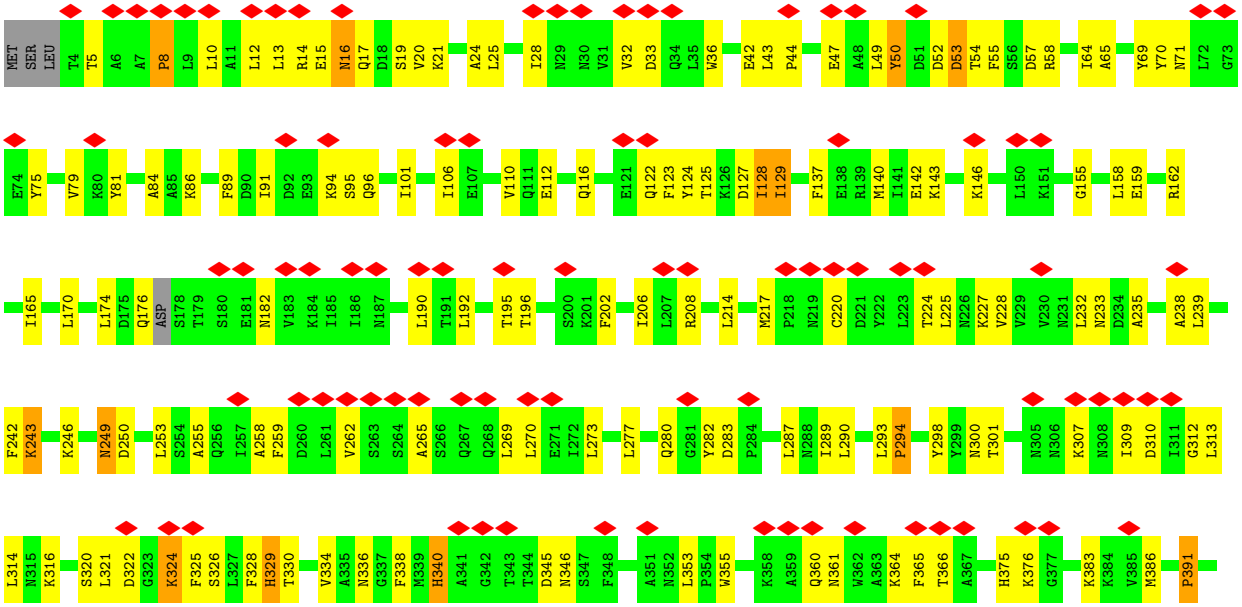
- Molecule 22: 26S PROTEASE REGULATORY SUBUNIT 6A

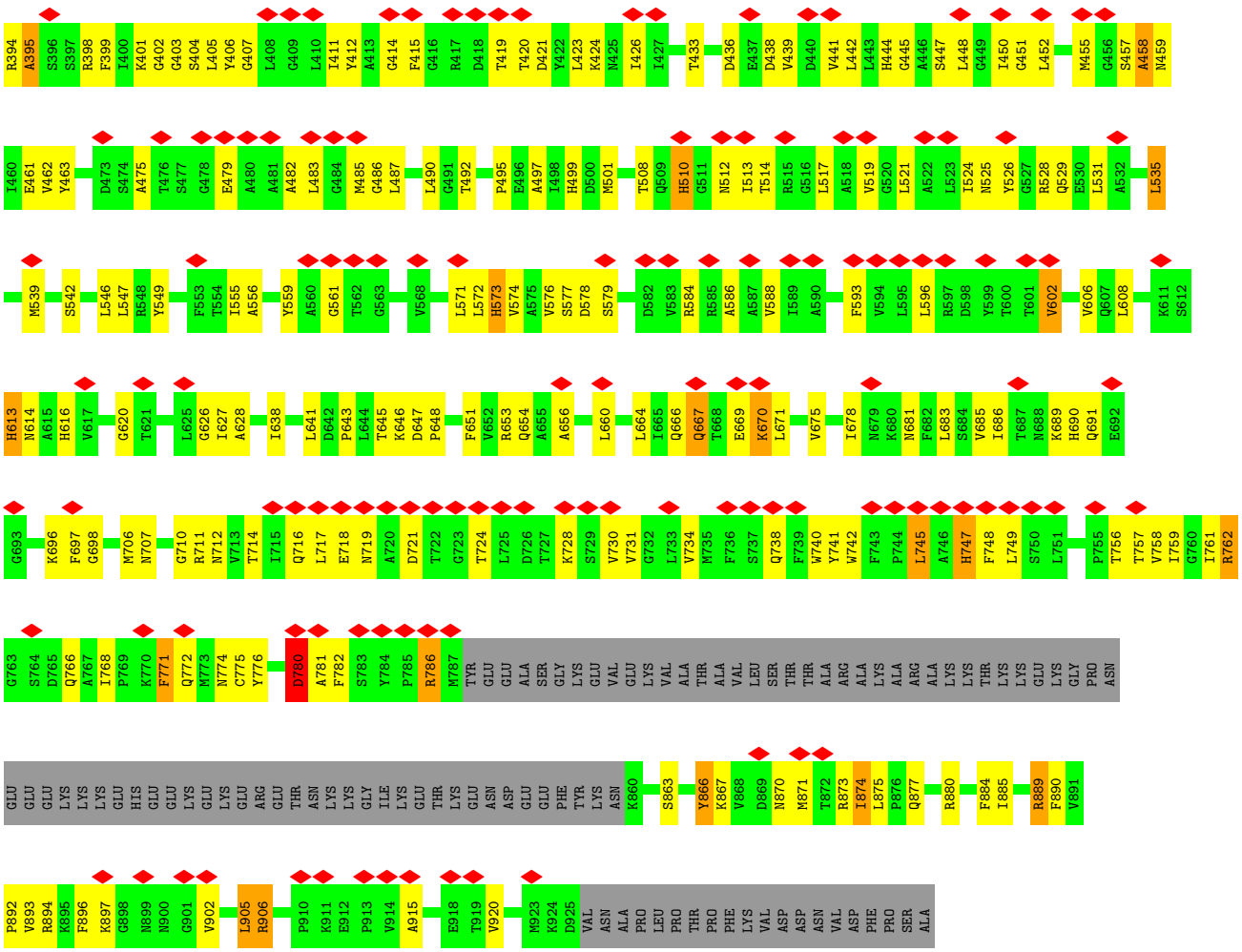
Chain M:



- Molecule 23: 26S PROTEASOME REGULATORY SUBUNIT RPN2

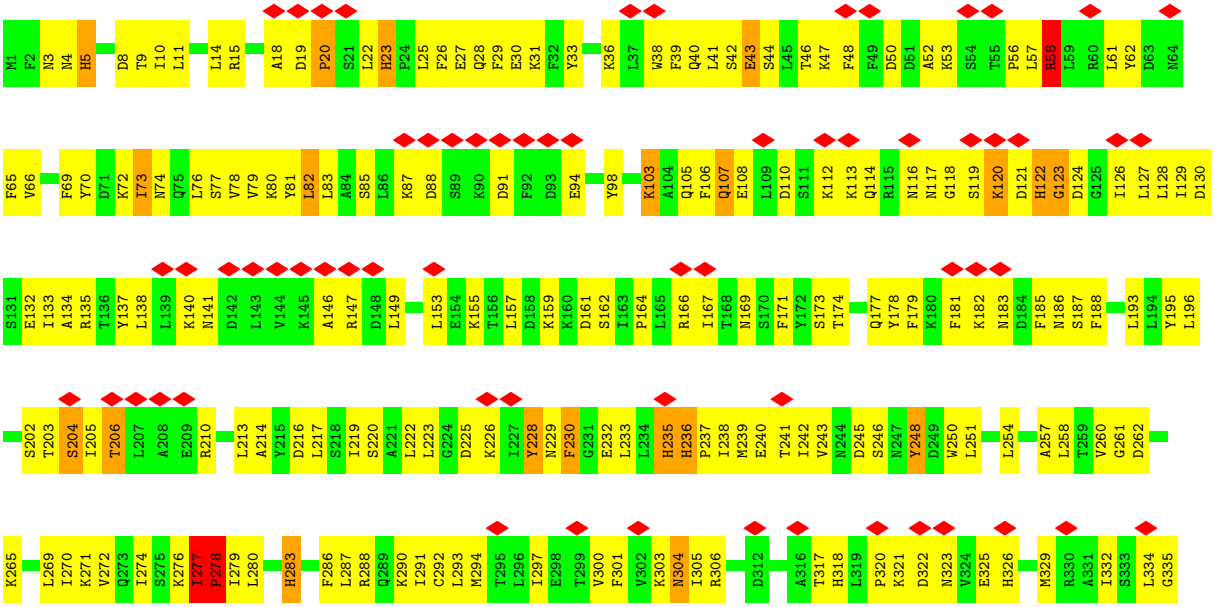
Chain N:

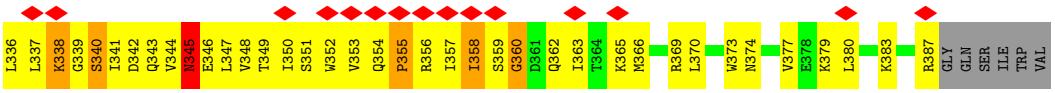




- Molecule 24: 26S PROTEASOME REGULATORY SUBUNIT RPN9

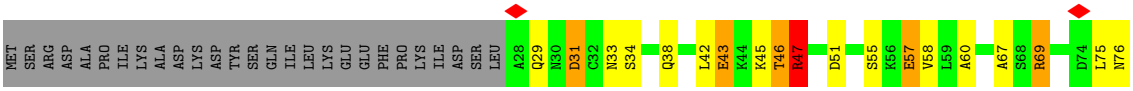
Chain O:

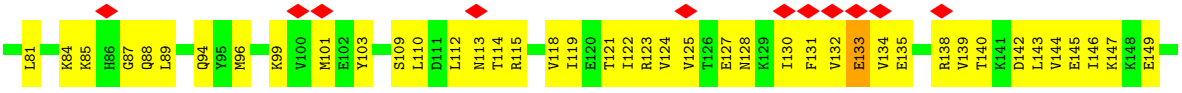


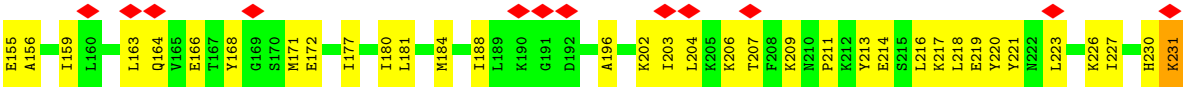


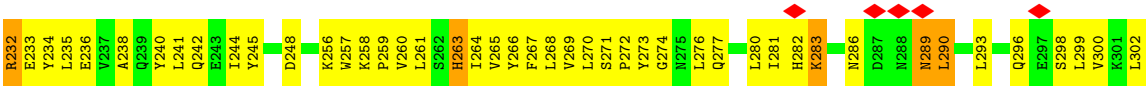
• Molecule 25: 26S PROTEASOME REGULATORY SUBUNIT RPN5

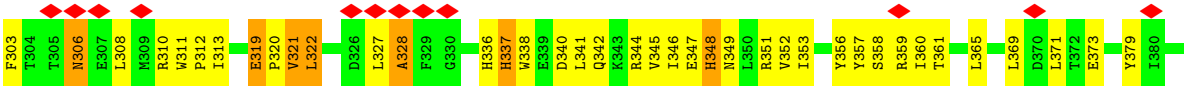
Chain P:

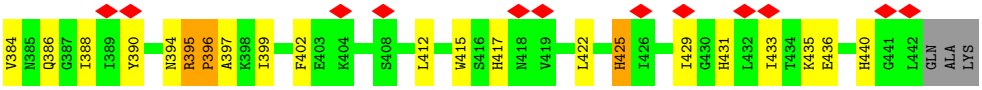






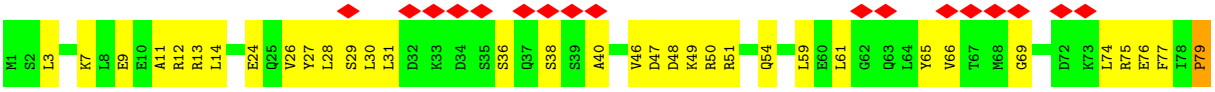


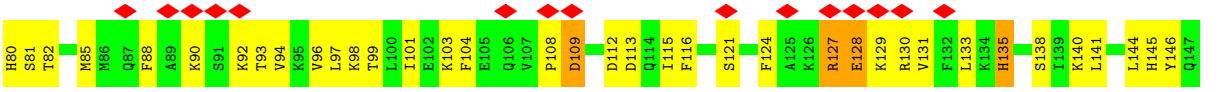


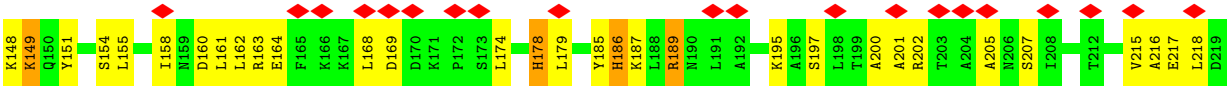


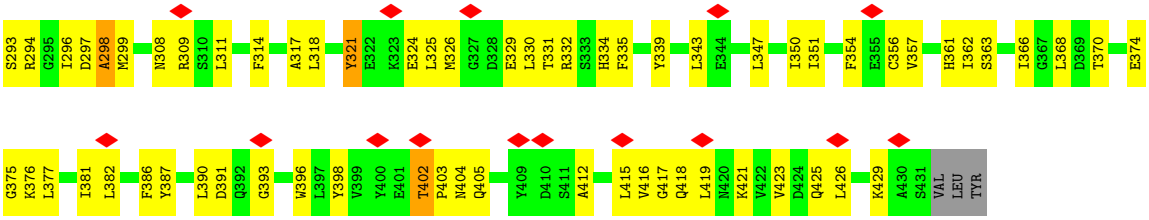
• Molecule 26: 26S PROTEASOME REGULATORY SUBUNIT RPN6

Chain Q:



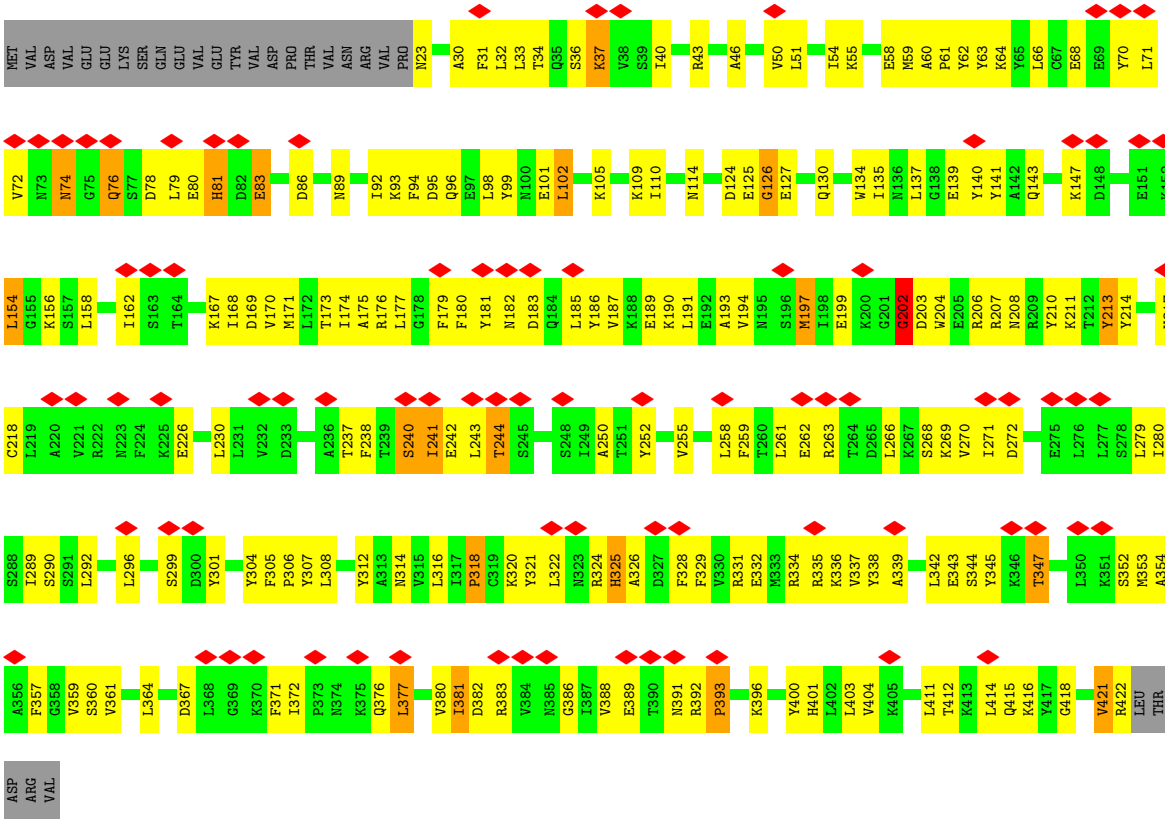






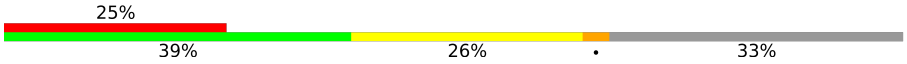
• Molecule 27: 26S PROTEASOME REGULATORY SUBUNIT RPN7

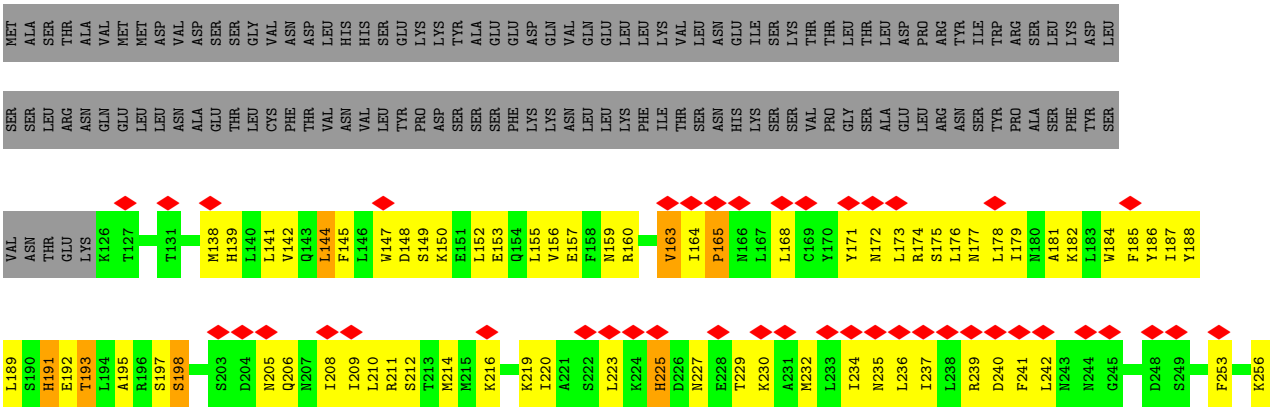
Chain R:

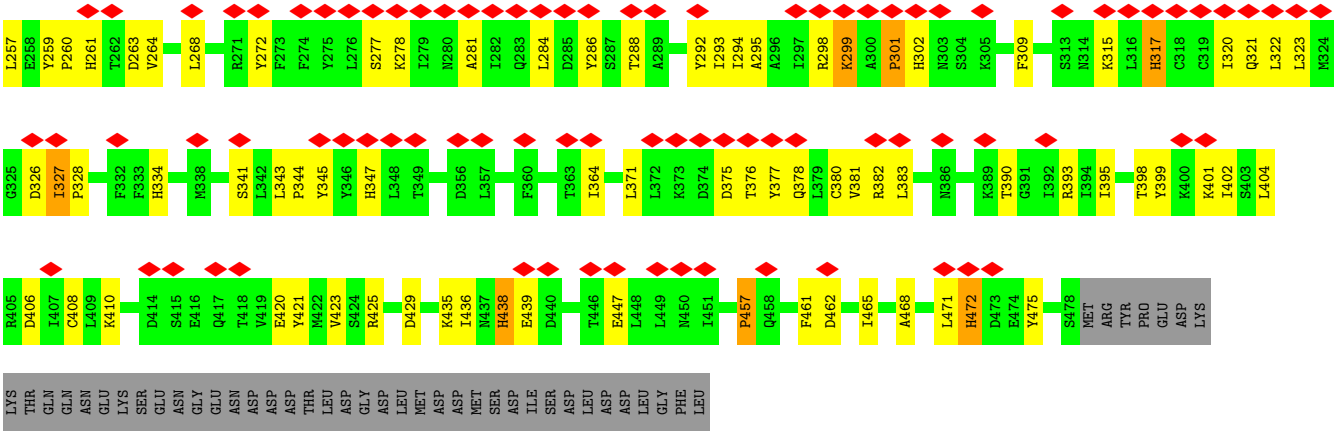


• Molecule 28: 26S PROTEASOME REGULATORY SUBUNIT RPN3

Chain S:

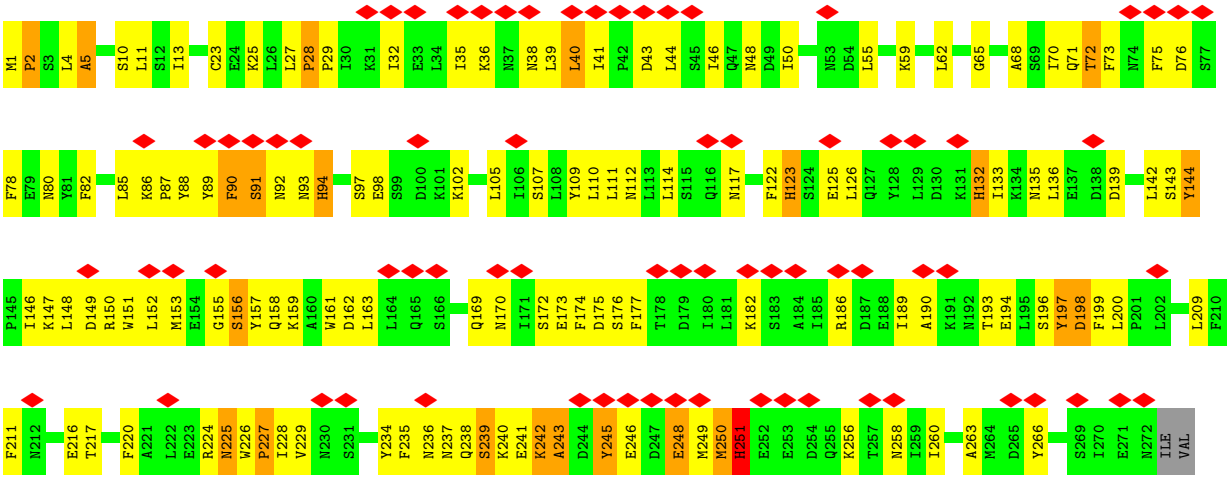






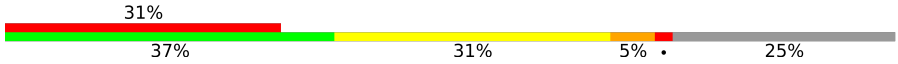
• Molecule 29: 26S PROTEASOME REGULATORY SUBUNIT RPN12

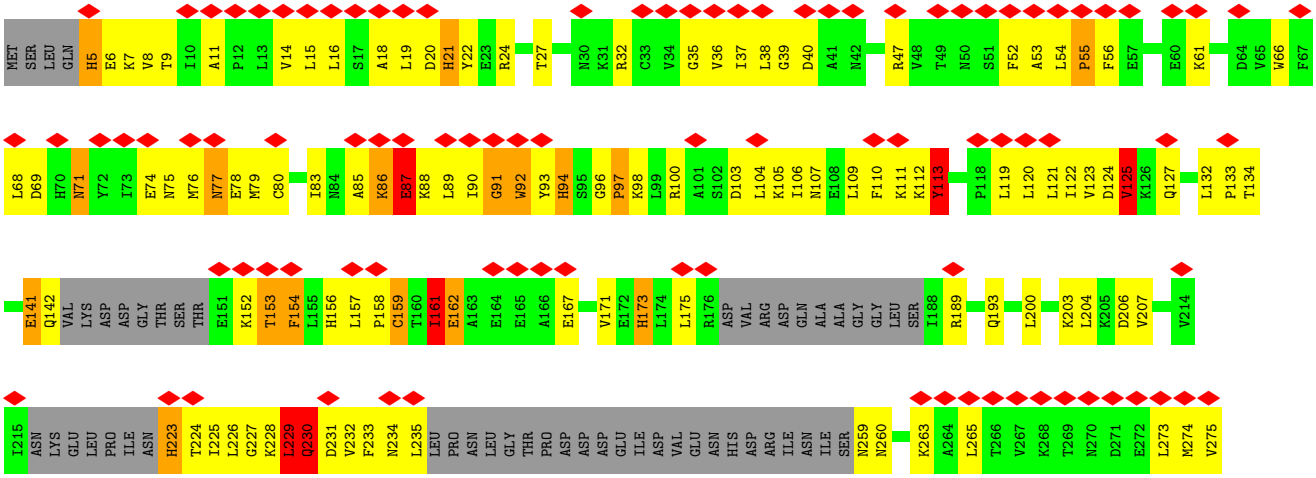
Chain T:

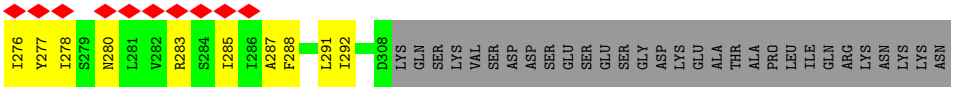


• Molecule 30: 26S PROTEASOME REGULATORY SUBUNIT RPN8

Chain U:

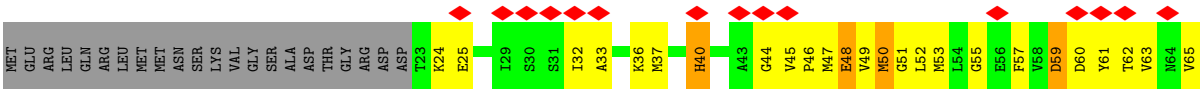


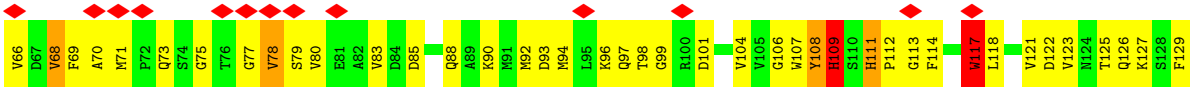


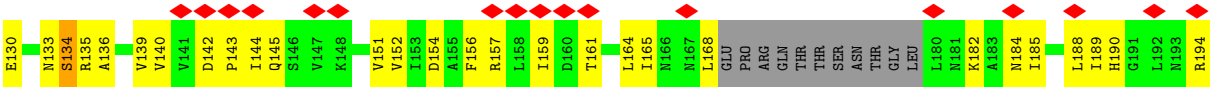


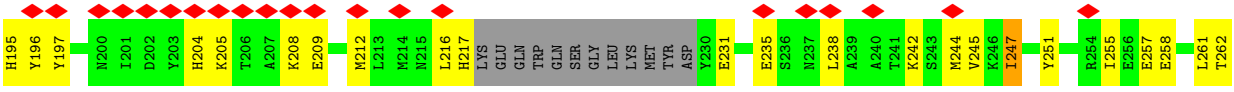
• Molecule 31: 26S PROTEASOME REGULATORY SUBUNIT RPN11

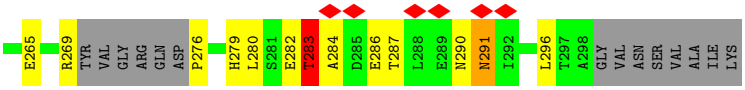
Chain V:





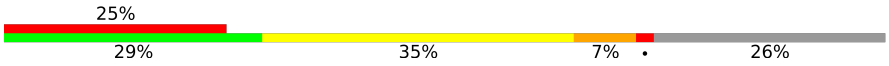


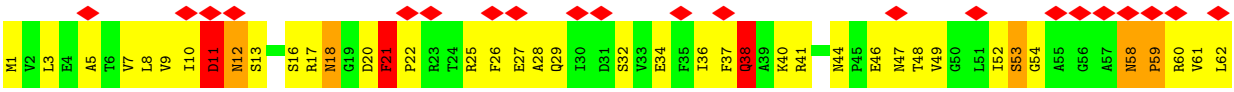


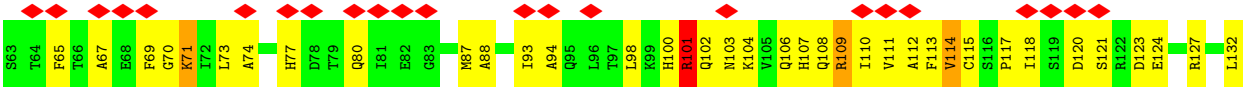


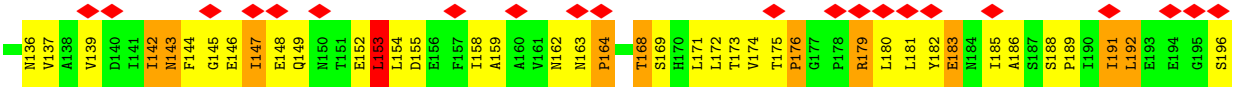
• Molecule 32: 26S PROTEASOME REGULATORY SUBUNIT RPN10

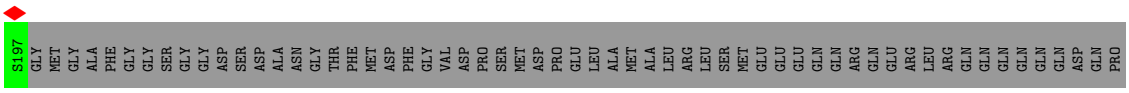
Chain W:





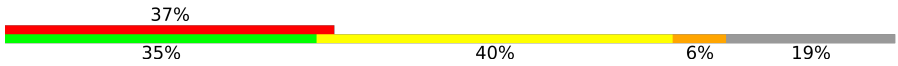


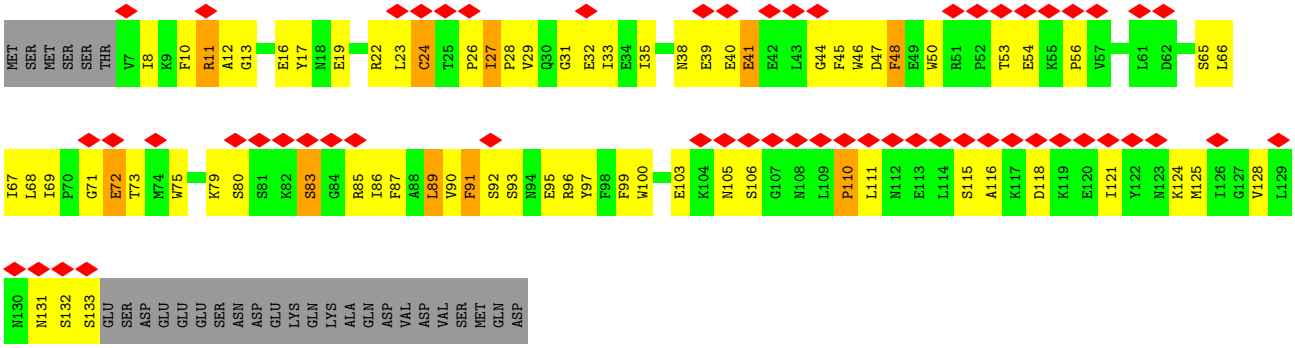




• Molecule 33: 26S PROTEASOME REGULATORY SUBUNIT RPN13

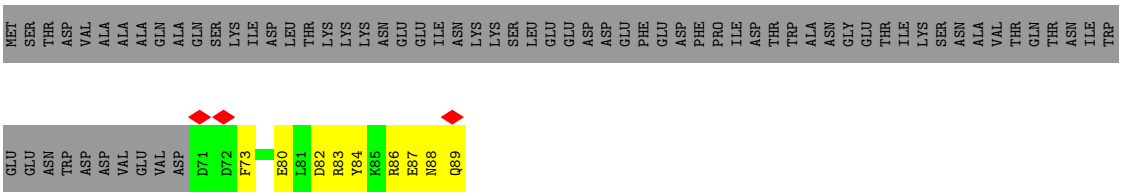
Chain X:





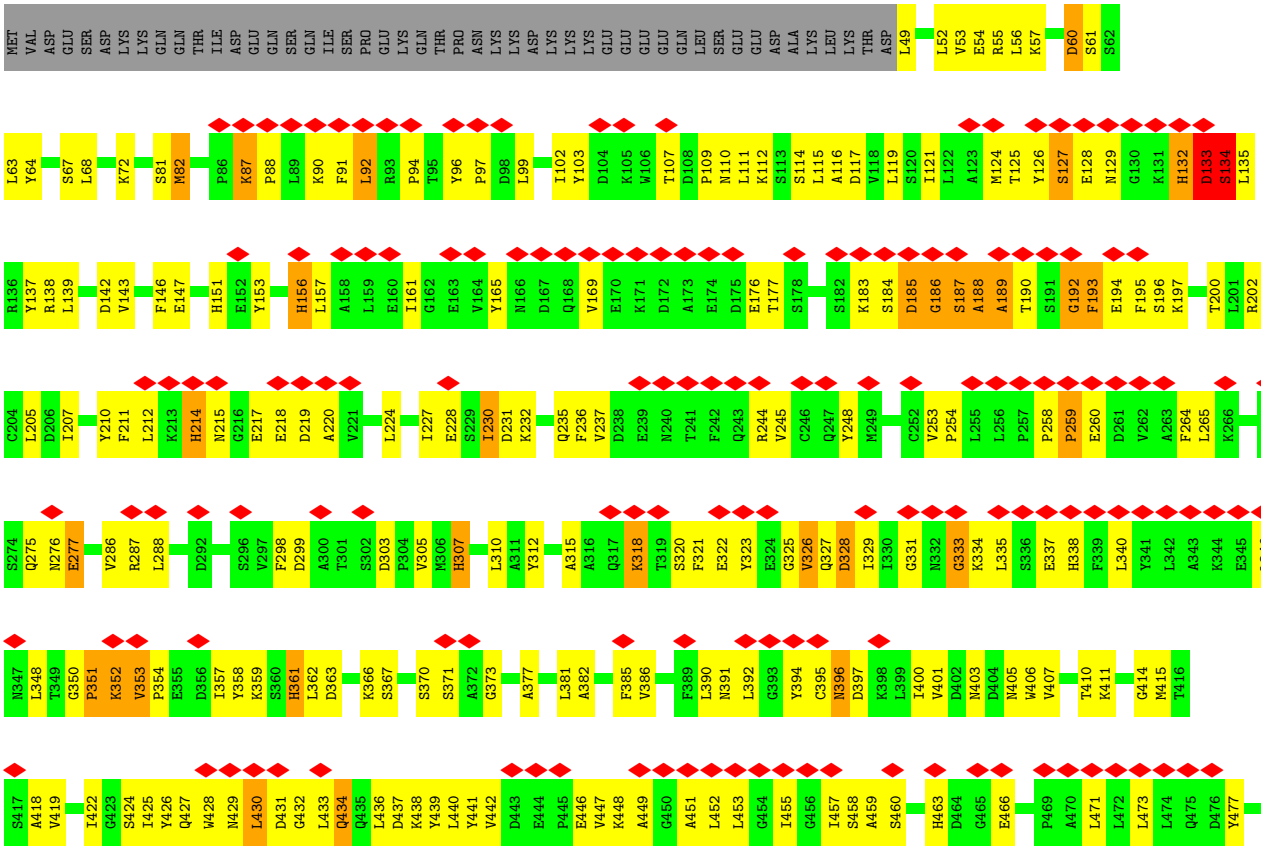
• Molecule 34: 26S PROTEASOME COMPLEX SUBUNIT SEM1

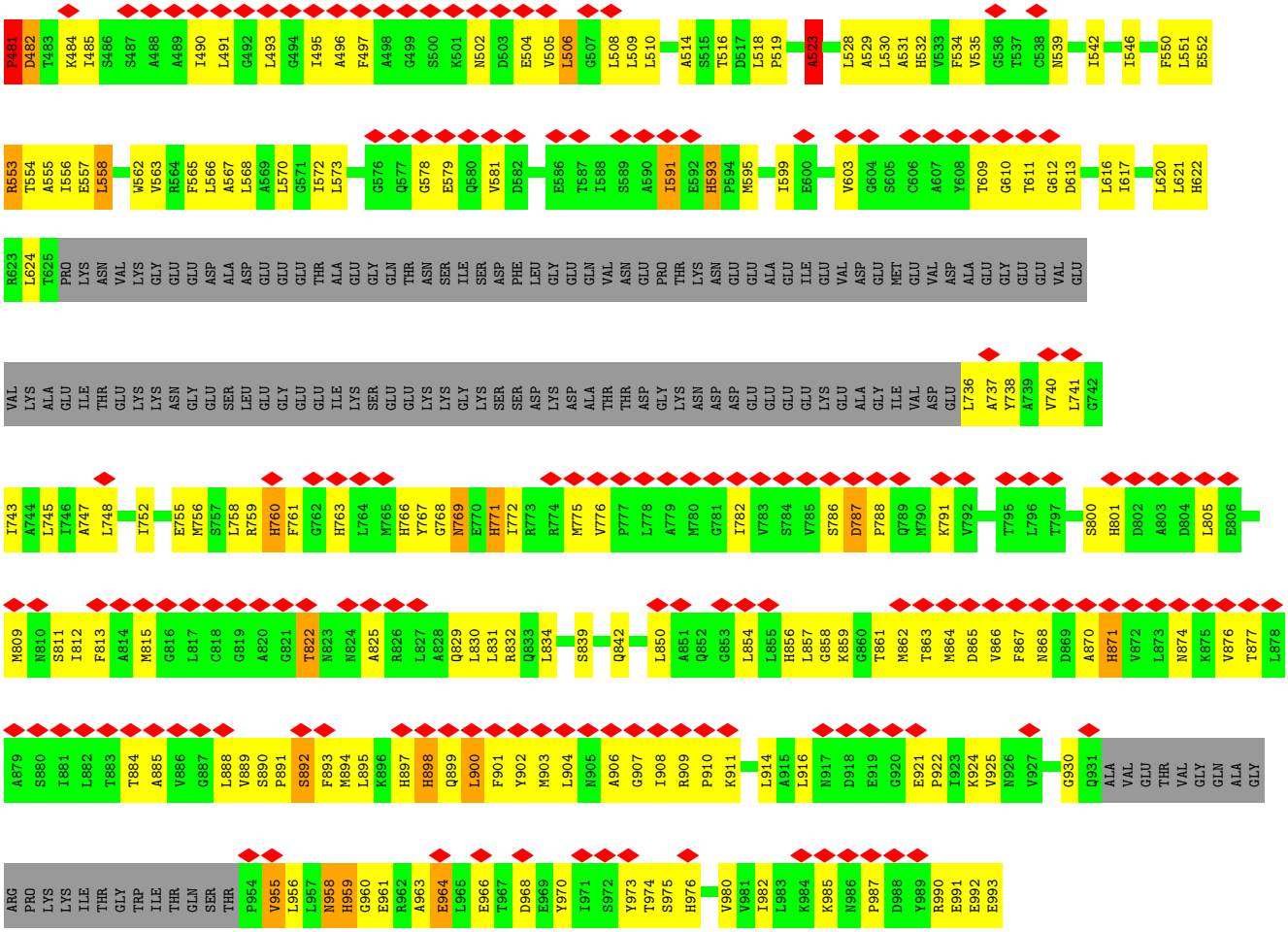
Chain Y:



• Molecule 35: 26S PROTEASOME REGULATORY SUBUNIT RPN1

Chain Z:




P481 D482 T483 K484 I485 S486 S487 A488 A489 I490 L491 G492 L493 G494 I495 A496 F497 A498 G499 S500 K501 N502 D503 E504 V505 L506 G507 L508 L509 L510 A514 S515 T516 D517 L518 P519 A523 L528 A529 L530 A531 H532 V533 F534 V535 G536 T537 C538 N539 I542 I546 F550 L551 E552
R553 T554 A555 I556 E557 L558 W562 V563 R564 F565 L566 A567 L568 A569 L570 G571 T572 L573 G576 Q577 G578 E579 Q580 V581 D582 E586 T587 I588 S589 A590 I591 E592 H593 F594 M595 I599 E600 V603 G604 S605 C606 A607 Y608 T609 G610 T611 G612 D613 L616 I617 L620 L621 H622
R623 L624 T625 PRO LYS ASN VAL LYS GLY GLU GLU ASP ALA ASP GLU GLU GLU THR ALA GLU GLY GLN THR ASN SER ILE SER ASP PHE LEU GLY GLU GLN VAL ASN GLU PRO THR LYS ASN GLU GLU ALA GLU ILE GLU VAL ASP GLU MET GLU VAL ASP ALA GLU GLY GLU GLU VAL GLU
VAL LYS ALA GLU ILE THR GLU LYS LYS ASN GLY GLU SER LEU GLU GLY GLU GLU ILE LYS SER GLU GLU LYS LYS GLY LYS SER SER ASP LYS ASP ALA THR THR ASP GLY LYS ASN ASP ASP GLU GLU GLU GLU LYS GLU ALA GLY ILE VAL ASP GLU L736 A737 Y738 A739 V740 L741 G742
I743 A744 L745 I746 A747 L748 I752 E755 M756 S757 L758 R759 H760 F761 G762 H763 L764 M765 H766 Y767 G768 N769 E770 H771 I772 R773 R774 M775 V776 P777 L778 A779 M780 G781 I782 V783 S784 V785 S786 D787 P788 Q789 M790 K791 V792 T795 L796 T797 S800 H801 D802 A803 D804 L805 E806
M809 N810 S811 I812 F813 A814 M815 G816 L817 C818 G819 A820 G821 T822 N823 N824 A825 R826 L827 A828 Q829 L830 L831 R832 Q833 L834 S839 Q842 L850 A851 Q852 G853 L854 L855 H856 L857 G858 K859 G860 T861 M862 T863 M864 D865 V866 F867 N868 D869 A870 H871 V872 L873 N874 K875 V876 T877 L878
A879 S880 I881 L882 T883 T884 A885 V886 G887 L888 V889 S890 P891 S892 F893 M894 L895 K896 H897 H898 Q899 L900 F901 Y902 M903 L904 N905 A906 G907 I908 R909 P910 K911 L914 A915 L916 N917 D918 E919 G920 E921 P922 I923 K924 V925 N926 V927 G930 Q931 ALA VAL GLU THR VAL GLY GLN ALA GLY
ARG PRO LYS LYS ILE THR GLY TRP ILE THR GLN SER THR P954 V955 L956 L957 N958 H959 G960 E961 R962 A963 E964 L965 E966 T967 D968 E969 Y970 I971 S972 Y973 T974 S975 H976 V980 V981 I982 L983 K984 K985 N986 P987 D988 Y989 R990 E991 E992 E993

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	53000	Depositor
Resolution determination method	FSC	Depositor
CTF correction method	MICROGRAPH	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	45	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON II (4k x 4k)	Depositor
Maximum map value	12.581	Depositor
Minimum map value	-11.710	Depositor
Average map value	0.006	Depositor
Map value standard deviation	0.297	Depositor
Recommended contour level	1.17	Depositor
Map size (Å)	557.2, 557.2, 557.2	wwPDB
Map dimensions	280, 280, 280	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.99, 1.99, 1.99	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

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

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	1	1.21	8/1605 (0.5%)	1.20	5/2171 (0.2%)
2	2	1.24	14/1723 (0.8%)	1.25	4/2337 (0.2%)
3	3	1.18	7/1611 (0.4%)	1.17	2/2174 (0.1%)
4	4	1.16	8/1613 (0.5%)	1.22	2/2173 (0.1%)
5	5	1.19	8/1683 (0.5%)	1.22	3/2277 (0.1%)
6	6	1.22	11/1795 (0.6%)	1.18	4/2420 (0.2%)
7	7	1.14	6/1855 (0.3%)	1.14	2/2514 (0.1%)
8	8	0.64	2/3093 (0.1%)	0.75	1/4141 (0.0%)
9	9	0.47	0/603	0.75	1/811 (0.1%)
10	A	1.23	10/1959 (0.5%)	1.21	10/2652 (0.4%)
11	B	1.21	8/1952 (0.4%)	1.21	5/2642 (0.2%)
12	C	1.16	11/1943 (0.6%)	1.19	2/2629 (0.1%)
13	D	1.12	7/1928 (0.4%)	1.13	2/2610 (0.1%)
14	E	1.19	8/1892 (0.4%)	1.14	2/2549 (0.1%)
15	F	1.23	13/1823 (0.7%)	1.16	4/2463 (0.2%)
16	G	1.24	15/1940 (0.8%)	1.22	4/2619 (0.2%)
17	H	1.08	7/2831 (0.2%)	1.28	11/3808 (0.3%)
18	I	1.17	14/2859 (0.5%)	1.20	8/3853 (0.2%)
19	J	1.17	13/2962 (0.4%)	1.15	4/3975 (0.1%)
20	K	1.21	10/3061 (0.3%)	1.37	11/4129 (0.3%)
21	L	1.16	11/2895 (0.4%)	1.14	3/3892 (0.1%)
22	M	1.17	10/2903 (0.3%)	1.24	14/3909 (0.4%)
23	N	1.15	31/6670 (0.5%)	1.21	26/9023 (0.3%)
24	O	0.73	10/3243 (0.3%)	0.95	4/4374 (0.1%)
25	P	1.17	12/3452 (0.3%)	1.20	12/4657 (0.3%)
26	Q	1.11	16/3527 (0.5%)	1.09	7/4748 (0.1%)
27	R	1.05	12/3272 (0.4%)	1.08	4/4412 (0.1%)
28	S	1.08	13/2945 (0.4%)	1.11	2/3976 (0.1%)
29	T	1.04	9/2279 (0.4%)	1.07	3/3077 (0.1%)
30	U	1.82	12/2087 (0.6%)	1.16	11/2811 (0.4%)
31	V	1.20	12/1969 (0.6%)	1.26	16/2652 (0.6%)
32	W	1.40	12/1556 (0.8%)	1.73	16/2108 (0.8%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
33	X	1.20	4/1058 (0.4%)	1.33	5/1432 (0.3%)
34	Y	0.97	0/169	0.94	0/223
35	Z	1.03	31/6403 (0.5%)	1.12	15/8686 (0.2%)
All	All	1.14	375/85159 (0.4%)	1.17	225/114927 (0.2%)

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	1	0	1
6	6	0	1
7	7	0	2
10	A	0	3
13	D	0	1
15	F	0	1
17	H	0	6
19	J	0	1
21	L	0	4
22	M	0	2
23	N	0	7
24	O	0	2
25	P	0	3
26	Q	0	1
27	R	0	2
28	S	0	1
29	T	0	2
30	U	0	6
31	V	0	6
32	W	0	11
33	X	0	3
35	Z	0	7
All	All	0	73

The worst 5 of 375 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
30	U	230	GLN	N-CA	70.24	2.86	1.46
32	W	38	GLN	C-N	-38.10	0.46	1.34
20	K	242	PHE	C-N	21.53	1.83	1.34
22	M	257	GLY	C-N	20.41	1.80	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
35	Z	134	SER	N-CA	16.22	1.78	1.46

The worst 5 of 225 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
32	W	154	LEU	O-C-N	34.34	177.64	122.70
20	K	242	PHE	O-C-N	27.34	166.44	122.70
20	K	242	PHE	C-N-CA	-26.79	54.71	121.70
32	W	154	LEU	CA-C-N	-25.06	62.07	117.20
23	N	889	ARG	NE-CZ-NH1	23.10	131.85	120.30

There are no chirality outliers.

5 of 73 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1	-2	LEU	Peptide
6	6	-9	GLN	Peptide
7	7	100	ASN	Peptide
7	7	33	ARG	Peptide
10	A	46	ARG	Mainchain

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	1	1576	0	1555	199	0
2	2	1692	0	1699	93	0
3	3	1581	0	1574	102	0
4	4	1585	0	1590	91	0
5	5	1646	0	1595	122	0
6	6	1757	0	1711	86	0
7	7	1824	0	1832	150	0
8	8	3008	0	2929	210	0
9	9	601	0	624	148	0
10	A	1921	0	1909	159	0
11	B	1915	0	1929	144	0
12	C	1913	0	1914	146	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	D	1899	0	1908	108	0
14	E	1867	0	1841	161	0
15	F	1795	0	1797	178	0
16	G	1900	0	1888	192	0
17	H	2792	0	2878	342	0
18	I	2822	0	2869	288	0
19	J	2928	0	3054	358	0
20	K	3019	0	3079	411	0
21	L	2853	0	2925	284	0
22	M	2866	0	2936	288	0
23	N	6562	0	6625	481	0
24	O	3182	0	3207	604	0
25	P	3401	0	3483	321	0
26	Q	3471	0	3494	359	0
27	R	3218	0	3216	408	0
28	S	2893	0	2937	244	0
29	T	2235	0	2206	305	0
30	U	2061	0	2116	376	0
31	V	1942	0	1954	250	0
32	W	1534	0	1538	204	0
33	X	1032	0	1017	182	0
34	Y	168	0	153	17	0
35	Z	6289	0	6233	619	0
All	All	83748	0	84215	7435	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 44.

The worst 5 of 7435 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)
35:Z:321:PHE:CZ	35:Z:350:GLY:HA2	1.23	1.67
8:8:320:LYS:HE3	17:H:164:SER:CB	1.22	1.66
24:O:373:TRP:CD1	30:U:200:LEU:HD21	1.29	1.65
26:Q:243:PHE:HZ	26:Q:287:THR:CA	1.07	1.64
24:O:373:TRP:HD1	30:U:200:LEU:CD2	1.02	1.61

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	1	203/215 (94%)	189 (93%)	13 (6%)	1 (0%)	29	69
2	2	221/261 (85%)	207 (94%)	10 (4%)	4 (2%)	8	40
3	3	202/205 (98%)	188 (93%)	12 (6%)	2 (1%)	15	55
4	4	196/198 (99%)	181 (92%)	13 (7%)	2 (1%)	15	55
5	5	210/287 (73%)	194 (92%)	11 (5%)	5 (2%)	6	33
6	6	220/241 (91%)	206 (94%)	11 (5%)	3 (1%)	11	46
7	7	231/266 (87%)	208 (90%)	16 (7%)	7 (3%)	4	28
8	8	368/416 (88%)	364 (99%)	4 (1%)	0	100	100
9	9	74/76 (97%)	71 (96%)	3 (4%)	0	100	100
10	A	241/252 (96%)	217 (90%)	20 (8%)	4 (2%)	9	42
11	B	248/250 (99%)	227 (92%)	15 (6%)	6 (2%)	6	33
12	C	243/258 (94%)	224 (92%)	16 (7%)	3 (1%)	13	50
13	D	240/254 (94%)	221 (92%)	16 (7%)	3 (1%)	12	48
14	E	241/260 (93%)	225 (93%)	13 (5%)	3 (1%)	13	50
15	F	231/234 (99%)	209 (90%)	19 (8%)	3 (1%)	12	48
16	G	243/288 (84%)	224 (92%)	16 (7%)	3 (1%)	13	50
17	H	353/467 (76%)	298 (84%)	36 (10%)	19 (5%)	2	19
18	I	358/437 (82%)	326 (91%)	25 (7%)	7 (2%)	7	38
19	J	367/405 (91%)	337 (92%)	21 (6%)	9 (2%)	5	32
20	K	377/428 (88%)	336 (89%)	29 (8%)	12 (3%)	4	26
21	L	357/437 (82%)	326 (91%)	18 (5%)	13 (4%)	3	25
22	M	363/434 (84%)	322 (89%)	29 (8%)	12 (3%)	4	26
23	N	843/945 (89%)	786 (93%)	37 (4%)	20 (2%)	6	33
24	O	385/393 (98%)	309 (80%)	44 (11%)	32 (8%)	1	12
25	P	413/445 (93%)	374 (91%)	22 (5%)	17 (4%)	3	23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
26	Q	429/434 (99%)	390 (91%)	28 (6%)	11 (3%)	5	31
27	R	398/429 (93%)	345 (87%)	36 (9%)	17 (4%)	2	22
28	S	351/523 (67%)	308 (88%)	31 (9%)	12 (3%)	3	26
29	T	270/274 (98%)	235 (87%)	16 (6%)	19 (7%)	1	14
30	U	245/338 (72%)	226 (92%)	9 (4%)	10 (4%)	3	23
31	V	239/306 (78%)	212 (89%)	18 (8%)	9 (4%)	3	24
32	W	193/268 (72%)	156 (81%)	17 (9%)	20 (10%)	0	8
33	X	125/156 (80%)	101 (81%)	19 (15%)	5 (4%)	3	23
34	Y	17/89 (19%)	17 (100%)	0	0	100	100
35	Z	807/993 (81%)	690 (86%)	68 (8%)	49 (6%)	1	17
All	All	10502/12162 (86%)	9449 (90%)	711 (7%)	342 (3%)	6	26

5 of 342 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	4	150	ASP
11	B	203	GLU
13	D	31	THR
13	D	204	GLN
17	H	183	ILE

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	1	169/178 (95%)	169 (100%)	0	100	100
2	2	182/214 (85%)	181 (100%)	1 (0%)	88	93
3	3	172/173 (99%)	172 (100%)	0	100	100
4	4	175/175 (100%)	175 (100%)	0	100	100
5	5	169/235 (72%)	168 (99%)	1 (1%)	86	92
6	6	185/201 (92%)	185 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	7	199/224 (89%)	199 (100%)	0	100	100
8	8	331/369 (90%)	325 (98%)	6 (2%)	59	77
9	9	68/68 (100%)	64 (94%)	4 (6%)	19	45
10	A	207/210 (99%)	207 (100%)	0	100	100
11	B	209/209 (100%)	209 (100%)	0	100	100
12	C	204/216 (94%)	204 (100%)	0	100	100
13	D	214/226 (95%)	214 (100%)	0	100	100
14	E	199/215 (93%)	199 (100%)	0	100	100
15	F	192/193 (100%)	192 (100%)	0	100	100
16	G	201/239 (84%)	200 (100%)	1 (0%)	88	93
17	H	303/399 (76%)	302 (100%)	1 (0%)	92	95
18	I	319/385 (83%)	318 (100%)	1 (0%)	92	95
19	J	325/352 (92%)	324 (100%)	1 (0%)	92	95
20	K	334/374 (89%)	333 (100%)	1 (0%)	92	95
21	L	308/377 (82%)	308 (100%)	0	100	100
22	M	315/375 (84%)	315 (100%)	0	100	100
23	N	713/797 (90%)	710 (100%)	3 (0%)	91	94
24	O	363/368 (99%)	361 (99%)	2 (1%)	86	92
25	P	388/415 (94%)	380 (98%)	8 (2%)	53	72
26	Q	388/391 (99%)	388 (100%)	0	100	100
27	R	351/379 (93%)	351 (100%)	0	100	100
28	S	330/489 (68%)	330 (100%)	0	100	100
29	T	254/256 (99%)	254 (100%)	0	100	100
30	U	234/308 (76%)	228 (97%)	6 (3%)	46	66
31	V	217/268 (81%)	214 (99%)	3 (1%)	67	80
32	W	171/230 (74%)	169 (99%)	2 (1%)	71	83
33	X	116/144 (81%)	114 (98%)	2 (2%)	60	78
34	Y	18/81 (22%)	18 (100%)	0	100	100
35	Z	692/850 (81%)	692 (100%)	0	100	100
All	All	9215/10583 (87%)	9172 (100%)	43 (0%)	89	93

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
25	P	85	LYS
30	U	154	PHE
25	P	88	GLN
30	U	75	ASN
31	V	85	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 173 such sidechains are listed below:

Mol	Chain	Res	Type
25	P	76	ASN
29	T	225	ASN
25	P	315	GLN
27	R	114	ASN
30	U	259	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)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
9	GLZ	9	76	-	3,3,3	2.42	1 (33%)	0,2,2	-	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	GLZ	9	76	-	-	0/0/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	9	76	GLZ	O-C	4.14	1.43	1.19

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	9	76	GLZ	18	0

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
32	W	3
19	J	2
20	K	2
21	L	2
22	M	2
31	V	2
18	I	1

The worst 5 of 14 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	J	218:LEU	C	219:VAL	N	3.99
1	K	248:GLY	C	249:GLU	N	3.09
1	L	251:ILE	C	252:VAL	N	2.52
1	I	252:LEU	C	253:ILE	N	2.13
1	J	224:GLY	C	225:GLU	N	2.03

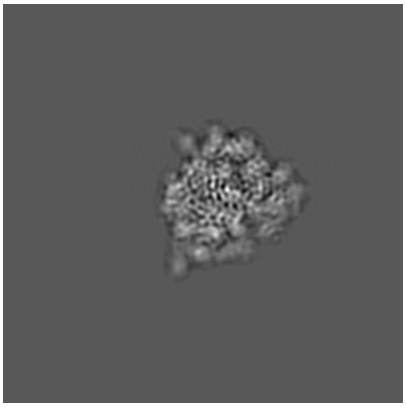
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-3034. These allow visual inspection of the internal detail of the map and identification of artifacts.

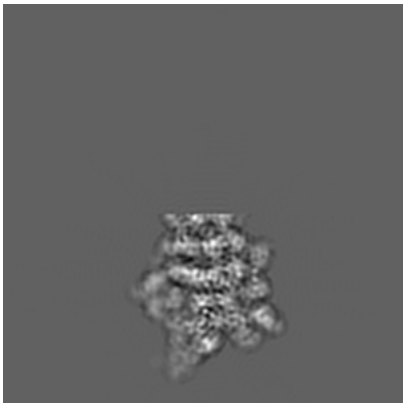
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

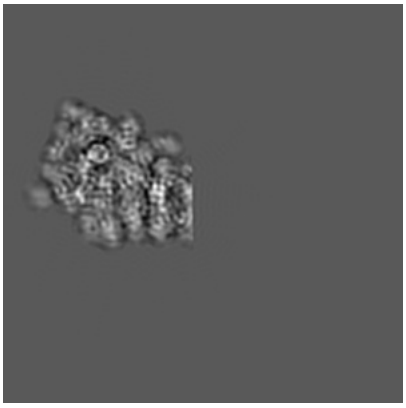
6.1.1 Primary map



X



Y

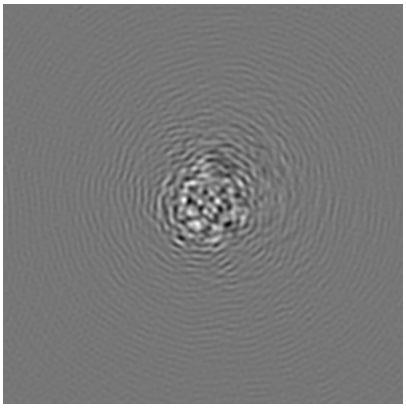


Z

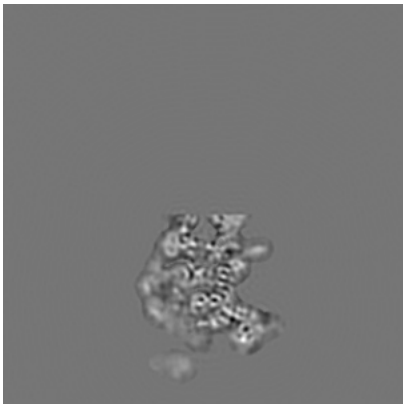
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

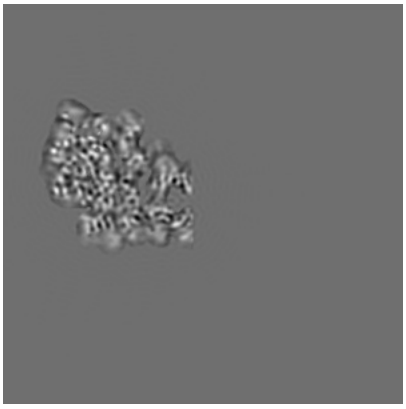
6.2.1 Primary map



X Index: 140



Y Index: 140



Z Index: 140

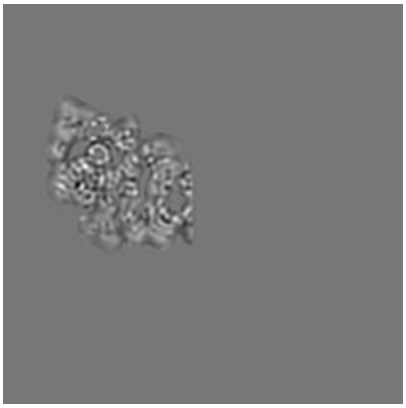
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map

X Index: 91

Y Index: 157

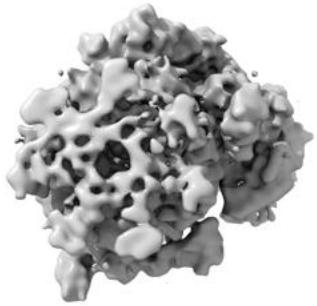


Z Index: 145

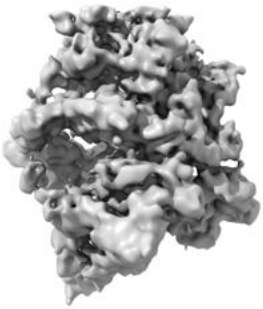
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

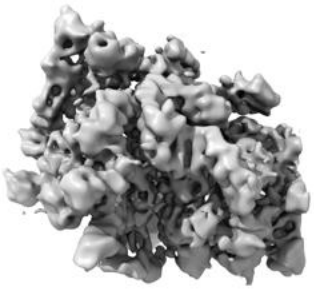
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 1.17. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

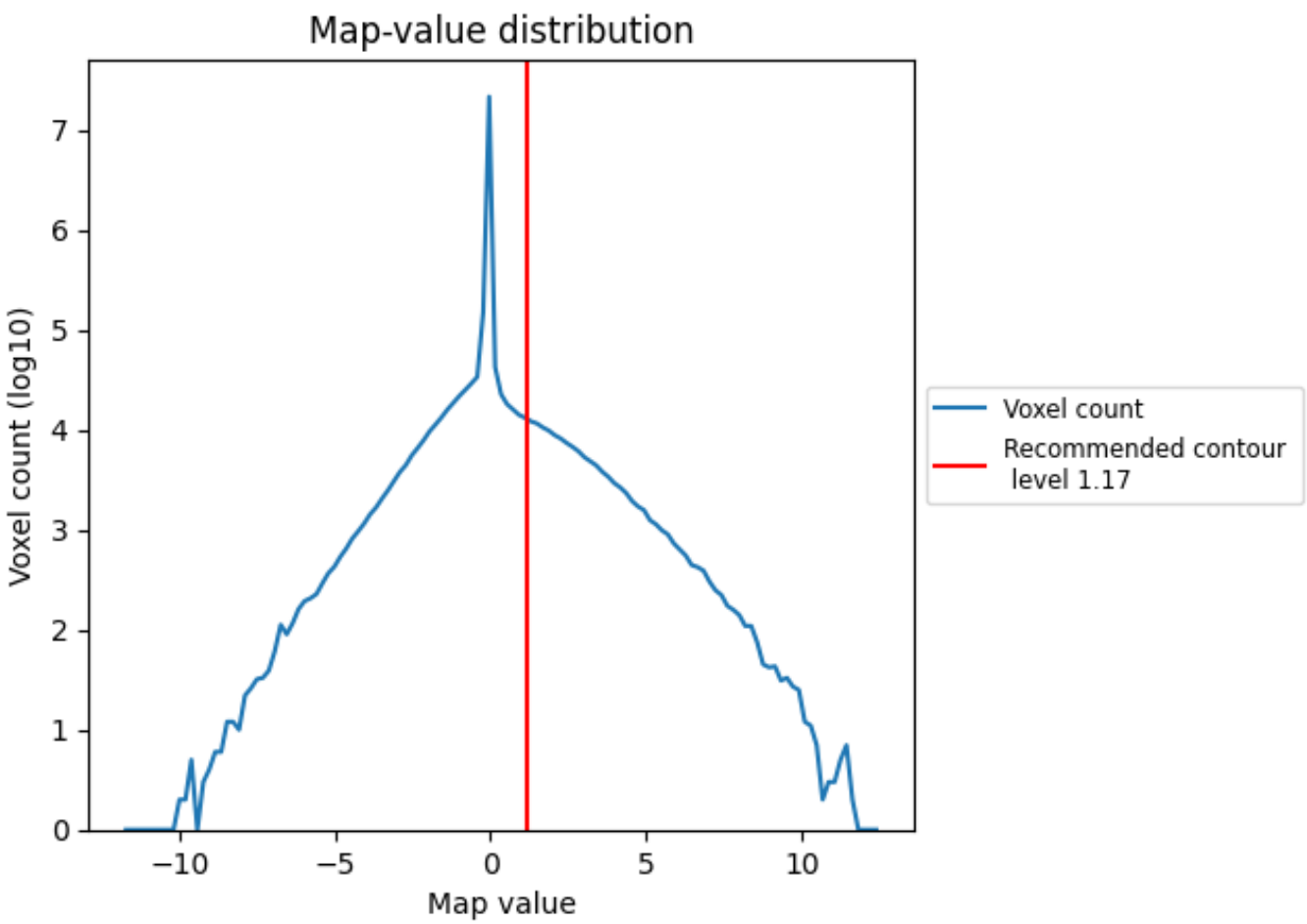
6.5 Mask visualisation (i)

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

7 Map analysis (i)

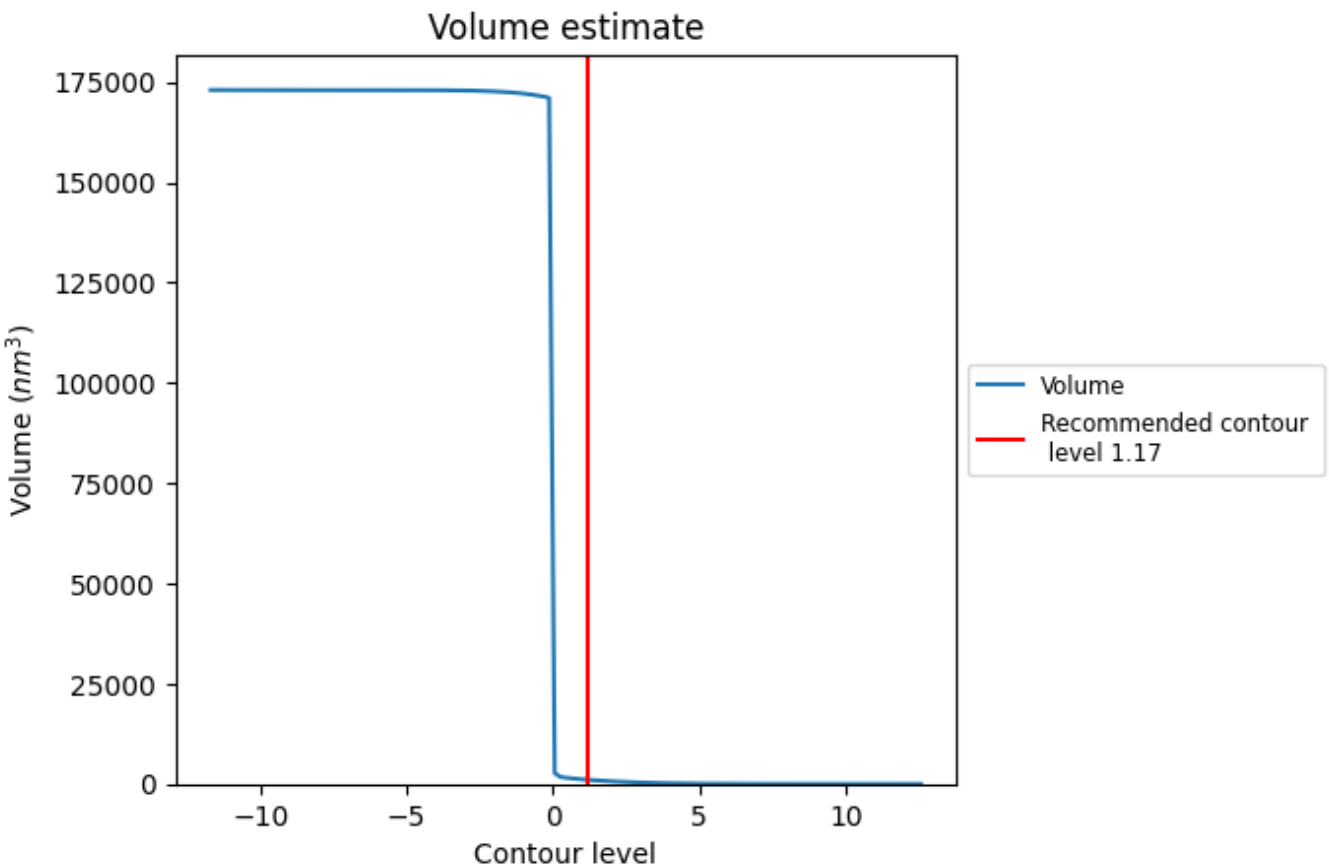
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



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

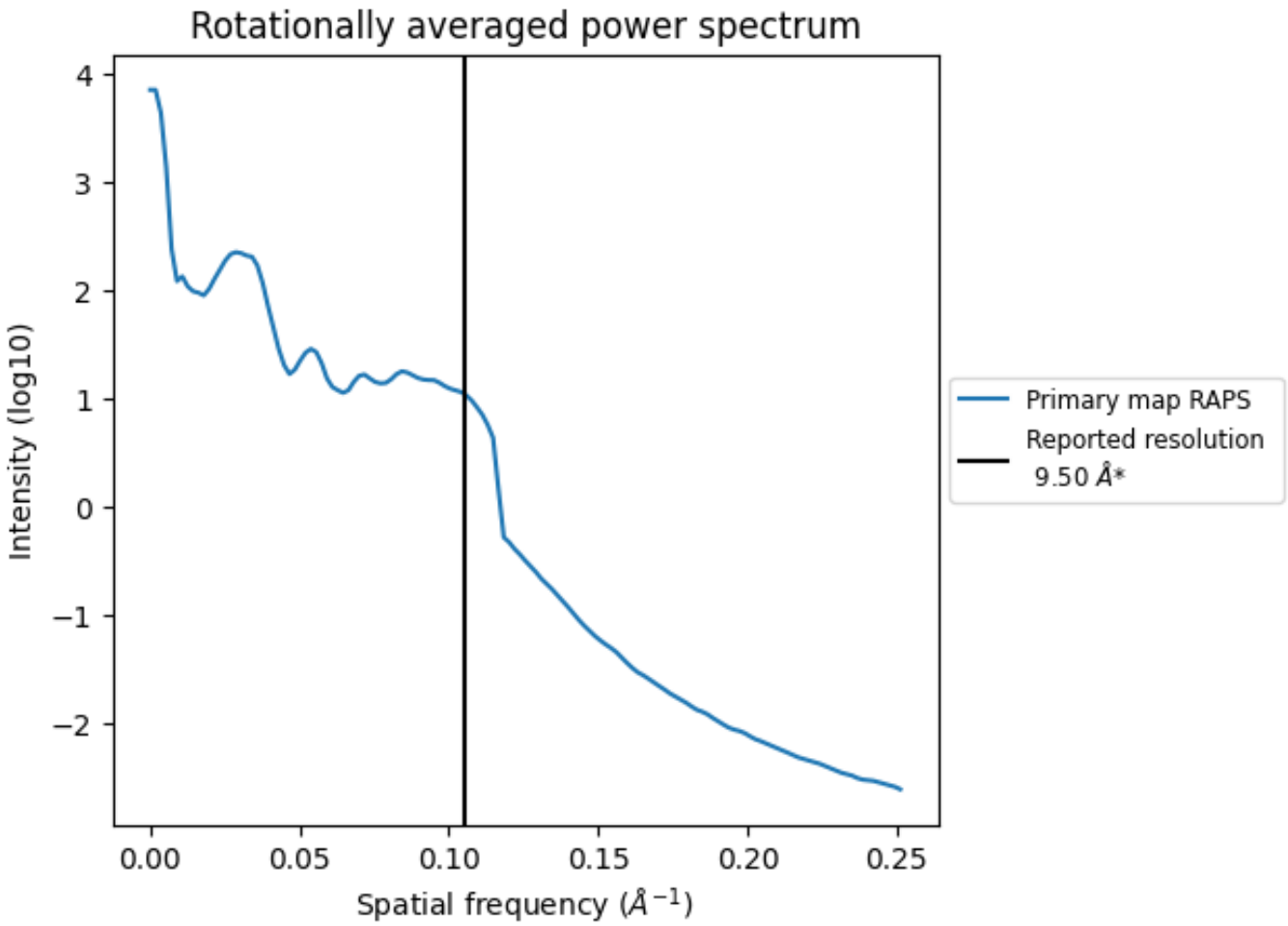
7.2 Volume estimate (i)



The volume at the recommended contour level is 1054 nm^3; this corresponds to an approximate mass of 952 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 0.105 Å^{-1}

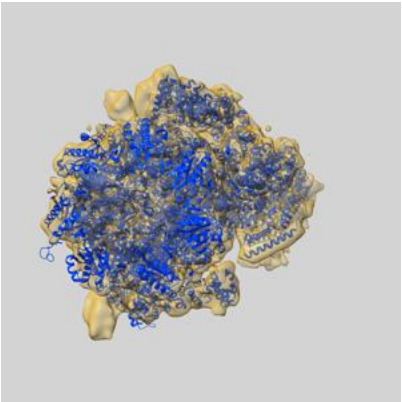
8 Fourier-Shell correlation (i)

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

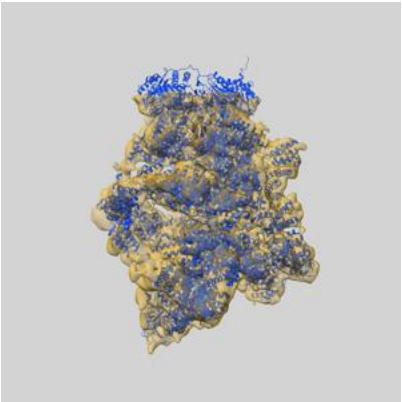
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-3034 and PDB model 5A5B. Per-residue inclusion information can be found in section 3 on page 11.

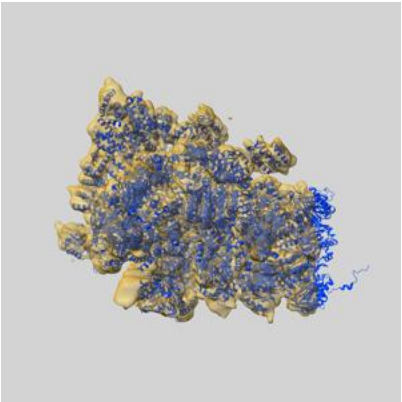
9.1 Map-model overlay (i)



X



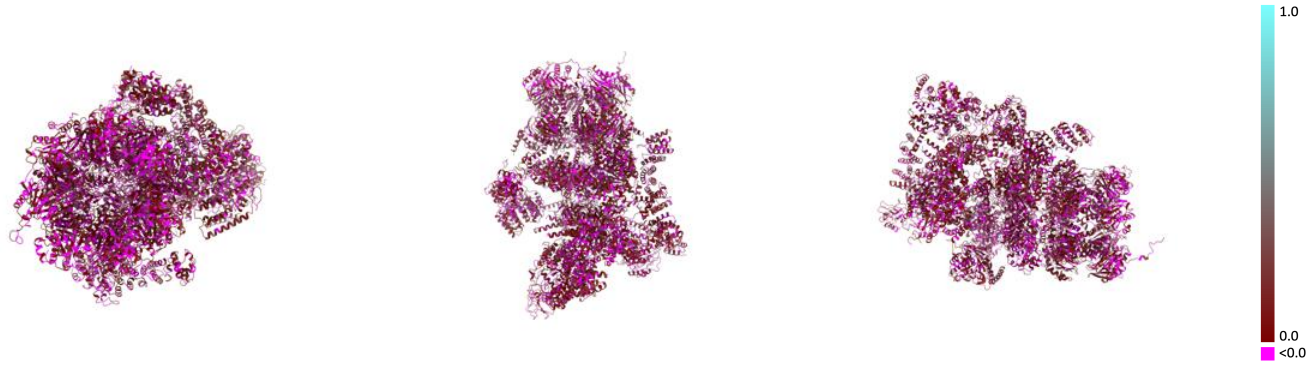
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 1.17 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.

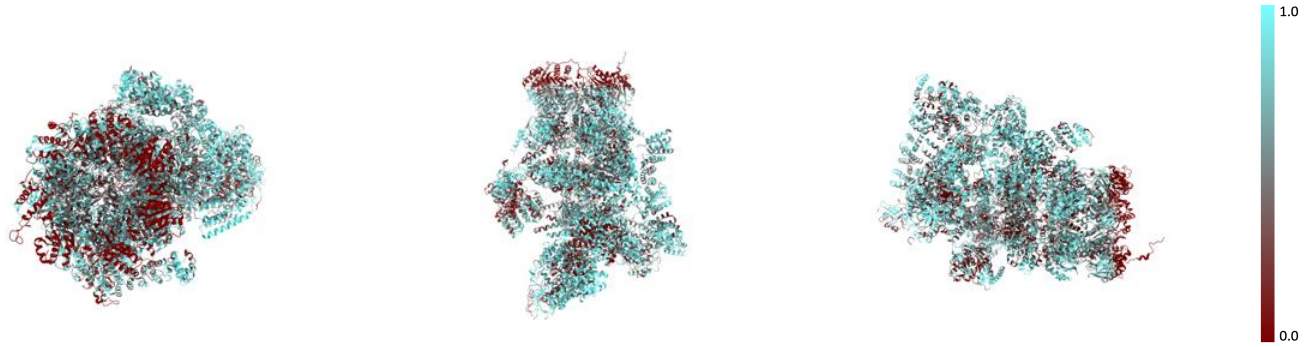
9.2 Q-score mapped to coordinate model (i)



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.

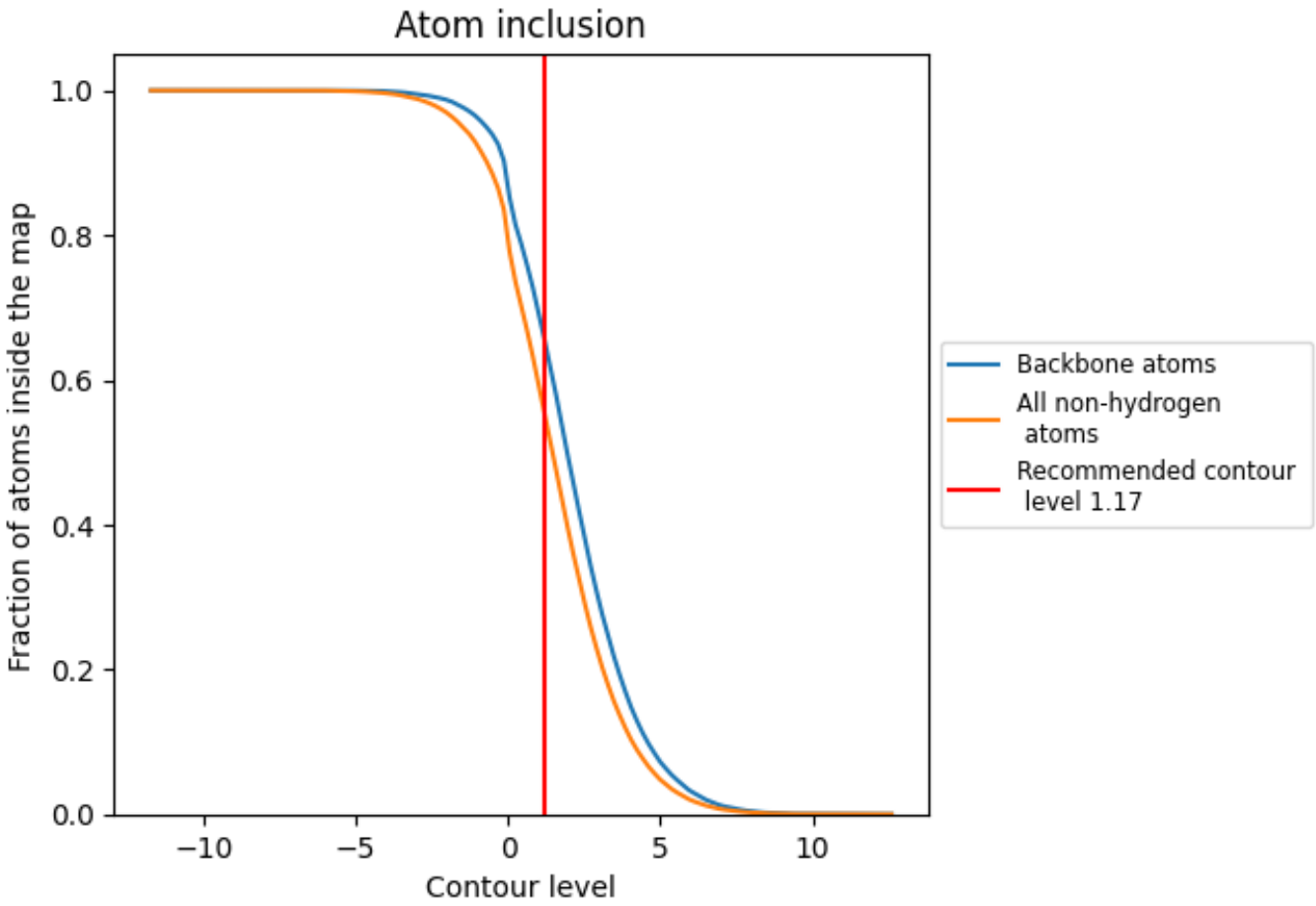
9.3 Atom inclusion mapped to coordinate model (i)



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 (1.17).

9.4 Atom inclusion (i)



At the recommended contour level, 66% of all backbone atoms, 56% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (1.17) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5601	0.0710
1	0.3508	0.0390
2	0.2591	0.0270
3	0.2564	0.0390
4	0.3595	0.0410
5	0.4015	0.0540
6	0.3902	0.0510
7	0.3956	0.0590
8	0.5409	0.0900
9	0.4476	0.0500
A	0.6180	0.0790
B	0.5533	0.0690
C	0.5694	0.0840
D	0.6602	0.1000
E	0.6270	0.0900
F	0.6600	0.0830
G	0.6241	0.0810
H	0.4620	0.0380
I	0.4986	0.0700
J	0.5640	0.0670
K	0.5200	0.0710
L	0.5680	0.0690
M	0.4952	0.0180
N	0.6472	0.0910
O	0.7098	0.0950
P	0.7463	0.0990
Q	0.7084	0.0990
R	0.6923	0.0860
S	0.5425	0.0710
T	0.6526	0.0800
U	0.4848	0.0760
V	0.5678	0.0830
W	0.5813	0.0500
X	0.5400	0.0430
Y	0.7362	0.0980
Z	0.5481	0.0520