W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB EM Validation Summary Report ⓘ

Dec 10, 2022 – 09:36 am GMT

PDB ID	:	5FMF
EMDB ID	:	EMD-3114
Title	:	the P-lobe of RNA polymerase II pre-initiation complex
Authors	:	Murakami, K.; Tsai, K.; Kalisman, N.; Bushnell, D.A.; Asturias, F.J.; Kornberg, R.D.
Deposited on	:	2015-11-03
Resolution	:	6.00 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

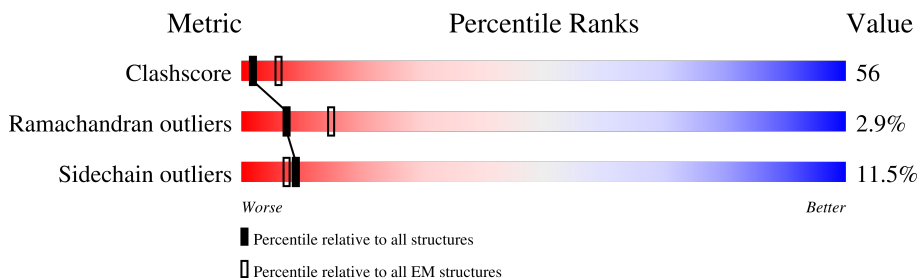|		
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 6.00 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	Percentile Ranks	Value
Clashscore		56
Ramachandran outliers		2.9%
Sidechain outliers		11.5%

Worse ← → *Better*

■ Percentile relative to all structures
▯ Percentile relative to all EM structures

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	492	66% / 35% 52% 8% · ·
2	2	174	60% / 64% 28% 6% ·
3	A	1733	37% / 53% 24% · · 18%
4	B	1224	52% / 61% 29% · 6%
5	C	266	34% / 73% 22% 5%
6	D	178	57% / 67% 26% 6% ·
7	E	214	43% / 74% 24% ·
8	F	84	27% / 60% 37% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	G	171	54% / 67% 27% 5%
10	H	146	35% / 57% 29% 5% 9%
11	I	119	52% / 76% 20%
12	J	65	32% / 57% 32% 11%
13	K	115	46% / 71% 28%
14	L	46	33% / 35% 41% 22%
15	M	116	63% / 69% 29%
16	N	72	56% / 94%
17	O	122	61% / 53% 25% 17%
18	P	345	74% / 9% 59% 13% 15%
19	Q	180	53% / 59% 36%
20	R	160	38% / 14% 46% 11% 26%
21	S	123	44% / 14% 49% 28% 7%
22	T	72	38% / 6% 92%
23	U	150	61% / 43% 51%
24	V	174	59% / 38% 55% 6%
25	W	62	92% / 90% 10%
26	X	63	83% / 90% 10%
27	Y	778	27% / 33% 38% 26%

2   Entry composition ⓘ

There are 29 unique types of molecules in this entry. The entry contains 55363 atoms, of which 691 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 DNA REPAIR HELICASE RAD25, SSL2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	1	473	Total	C	N	O	S	0	0
			3797	2390	664	720	23		

- Molecule 2 is a protein called TRANSCRIPTION ELONGATION FACTOR S-II, DST1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	2	174	Total	C	N	O	S	0	104
			666	454	99	108	5		

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	A	1422	Total	C	N	O	S	0	0
			11174	7036	1954	2122	62		

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	B	1156	Total	C	N	O	S	0	0
			9140	5781	1606	1697	56		

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	C	266	Total	C	N	O	S	0	0
			2095	1317	348	417	13		

- Molecule 6 is a protein called RNA POLYMERASE II PRE-INITIATION COMPLEX, RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	D	178	Total	C	N	O	S	0	0
			1434	887	257	288	2		

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC1, RPB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	E	214	Total 1752	C 1111	N 309	O 321	S 11	0	0

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC2, RPB6.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	F	84	Total 679	C 434	N 115	O 127	S 3	0	0

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	G	171	Total 1340	C 861	N 222	O 249	S 8	0	0

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC3, RPB8.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	H	133	Total 1068	C 673	N 180	O 211	S 4	0	0

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	I	119	Total 971	C 596	N 179	O 186	S 10	0	0

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC5, RPB10.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	J	65	Total 532	C 339	N 93	O 94	S 6	0	0

- Molecule 13 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	K	115	Total 920	C 590	N 157	O 171	S 2	0	1

- Molecule 14 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC4, RPB12.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	L	46	Total 363	C 224	N 72	O 63	S 4	0	0

- Molecule 15 is a protein called RNA POLYMERASE II PRE-INITIATION COMPLEX, TOA1.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	M	116	Total 956	C 599	N 159	O 195	S 3	0	0

- Molecule 16 is a DNA chain called NON-TEMPLATE STRAND DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	N	72	Total 1500	C 711	N 300	O 417	P 72	0	0

- Molecule 17 is a protein called TRANSCRIPTION INITIATION FACTOR IIA SUBUNIT 2, TOA2.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	O	101	Total 792	C 500	N 132	O 156	S 4	0	0

- Molecule 18 is a protein called TRANSCRIPTION INITIATION FACTOR IIB, SUA7.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	P	293	Total 2269	C 1437	N 386	O 428	S 18	0	0

- Molecule 19 is a protein called TATA-BOX-BINDING PROTEIN, TBP.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	Q	180	Total 1416	C 921	N 242	O 247	S 6	0	0

- Molecule 20 is a protein called TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT

ALPHA, TFA1.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	R	118	Total 960	C 625	N 158	O 172	S 5	0	0

- Molecule 21 is a protein called TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT BETA, TFA2.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	S	121	Total 1005	C 650	N 164	O 188	S 3	0	0

- Molecule 22 is a DNA chain called TEMPLATE STRAND DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	T	72	Total 1452	C 700	N 233	O 447	P 72	0	0

- Molecule 23 is a protein called RNA POLYMERASE II PRE-INITIATION COMPLEX, TFG1.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	U	150	Total 1248	C 793	N 220	O 232	S 3	0	0

- Molecule 24 is a protein called TRANSCRIPTION INITIATION FACTOR IIF SUBUNIT BETA, TFG2.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	V	174	Total 1455	C 930	N 250	O 269	S 6	0	0

- Molecule 25 is a protein called RNA POLYMERASE II TRANSCRIPTION FACTOR B SUBUNIT 2, TFB2.

Mol	Chain	Residues	Atoms				AltConf	Trace
25	W	62	Total 518	C 334	N 83	O 101	0	0

- Molecule 26 is a protein called RNA POLYMERASE II TRANSCRIPTION FACTOR B SUBUNIT 5, TFB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	X	63	Total	C	N	O	S	0	0
			499	316	88	93	2		

- Molecule 27 is a protein called DNA REPAIR HELICASE RAD3.

Mol	Chain	Residues	Atoms						AltConf	Trace
27	Y	578	Total	C	H	N	O	S	0	0
			5351	2975	691	790	862	33		

- Molecule 28 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
28	2	1	Total	Mg	0
			1	1	
28	A	1	Total	Mg	0
			1	1	

- Molecule 29 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
29	2	1	Total	Zn	0
			1	1	
29	A	2	Total	Zn	0
			2	2	
29	B	1	Total	Zn	0
			1	1	
29	C	1	Total	Zn	0
			1	1	
29	I	2	Total	Zn	0
			2	2	
29	J	1	Total	Zn	0
			1	1	
29	L	1	Total	Zn	0
			1	1	

3 Residue-property plots ⓘ

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: DNA REPAIR HELICASE RAD25, SSL2

Chain 1:
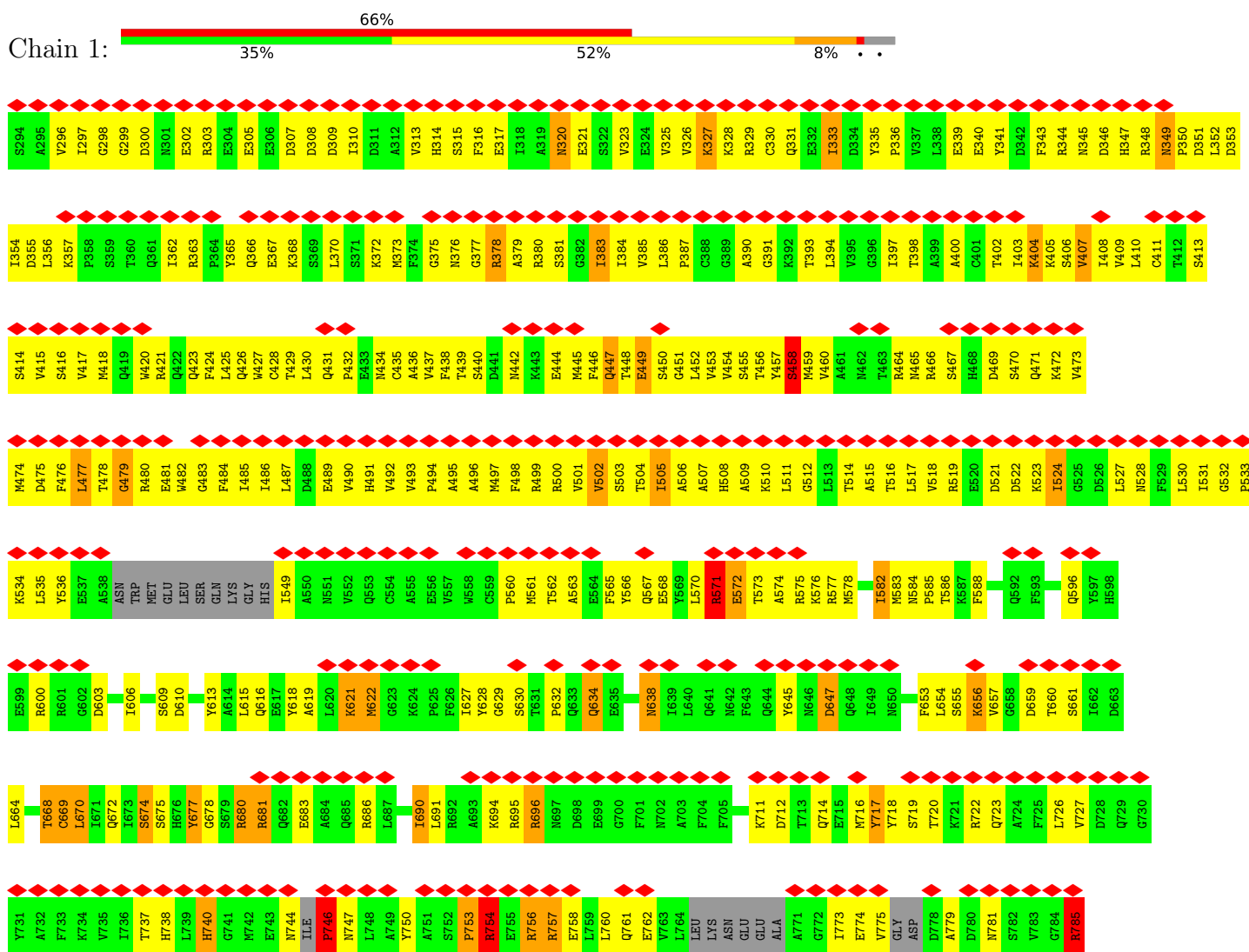
- Molecule 2: TRANSCRIPTION ELONGATION FACTOR S-II, DST1

Chain 2:
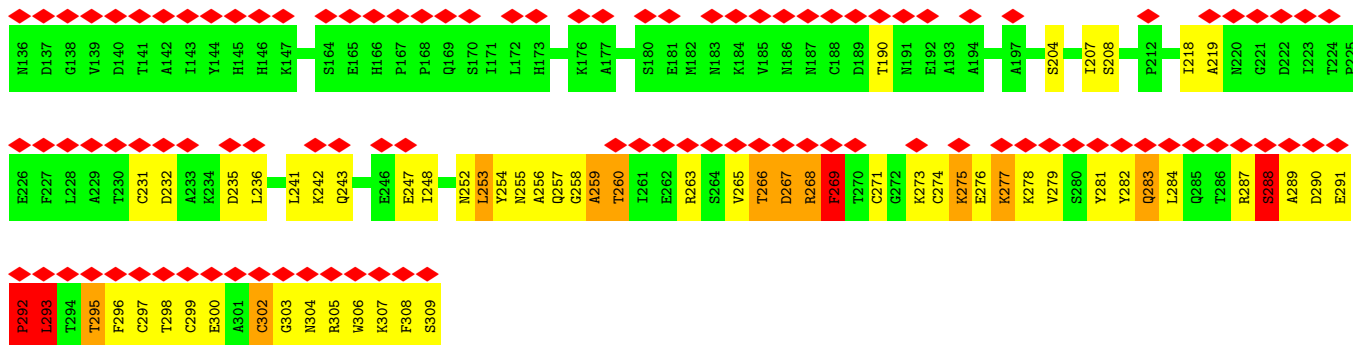
• Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1

Chain A:
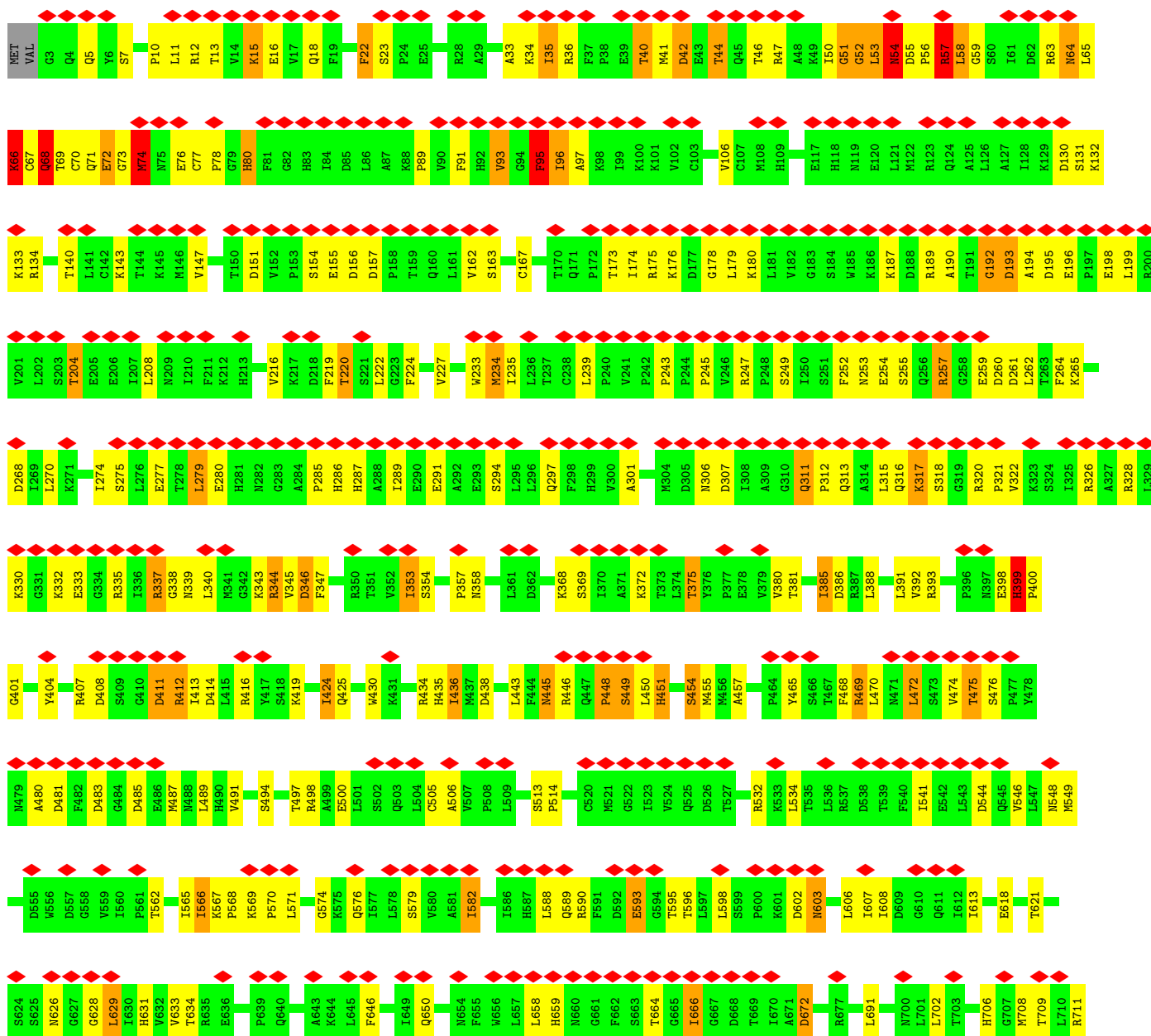
• Molecule 4: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2

Chain B:
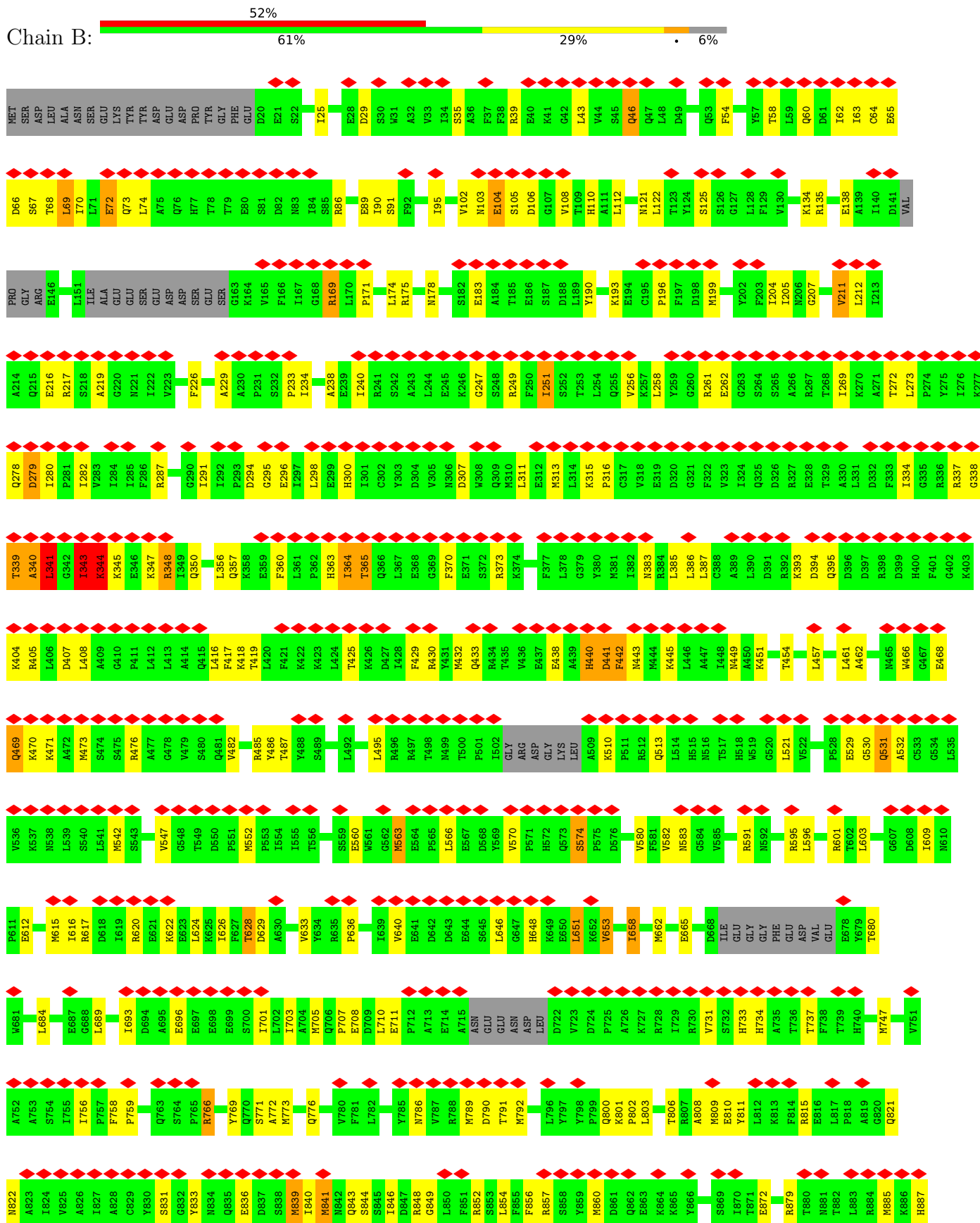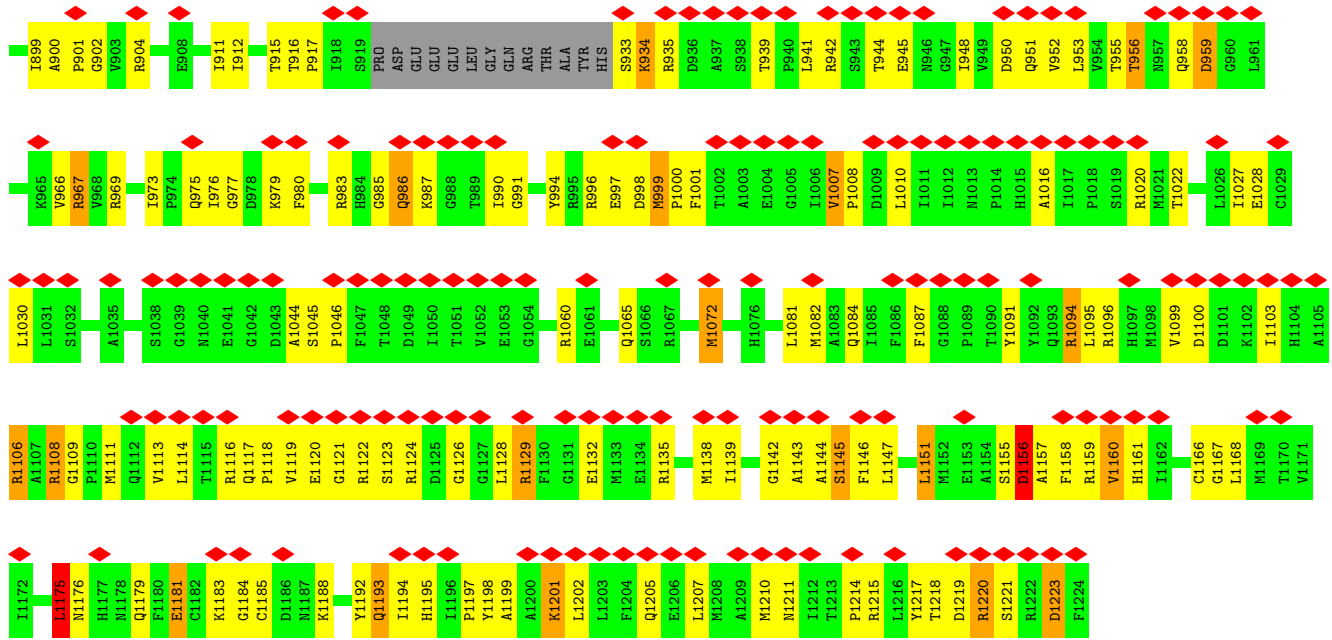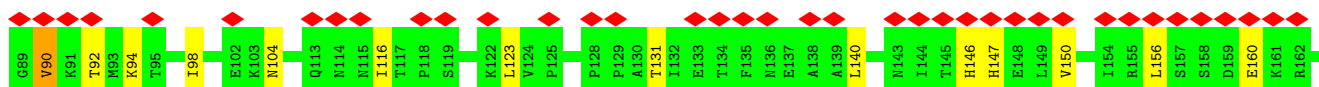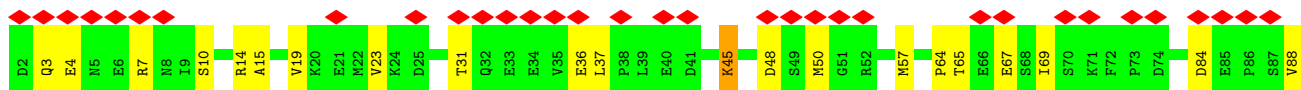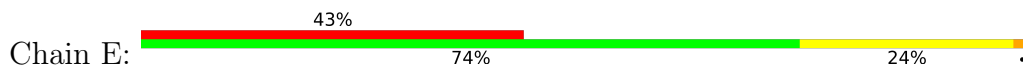
• Molecule 5: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3

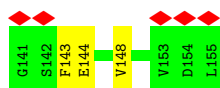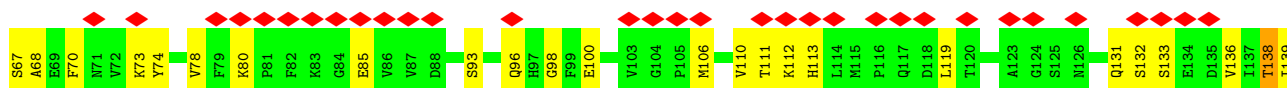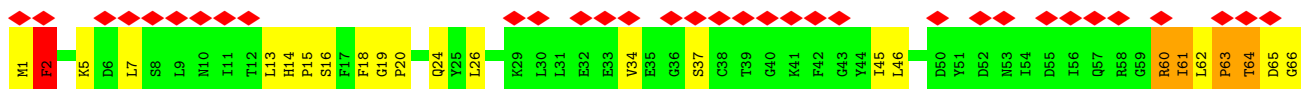Chain C:
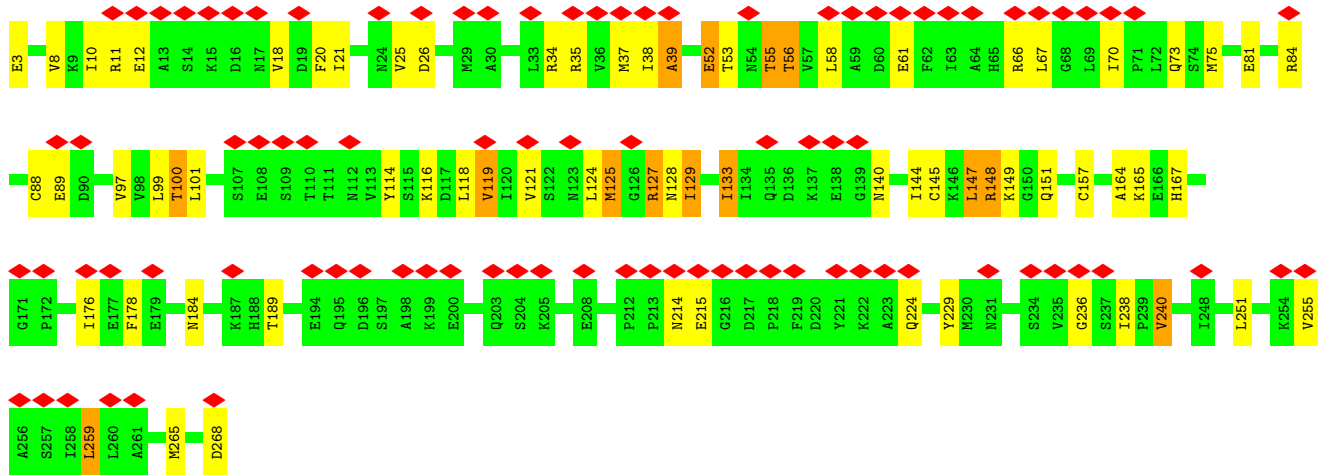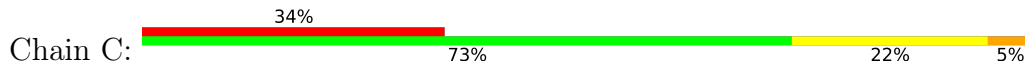
• Molecule 6: RNA POLYMERASE II PRE-INITIATION COMPLEX, RPB4

Chain D:

T187 T193 L194 I195 P196 S197 L198 N199 N200 K201 I202 S203 D204 D205 E206 R209 S215 N216 T219 L220 Y221
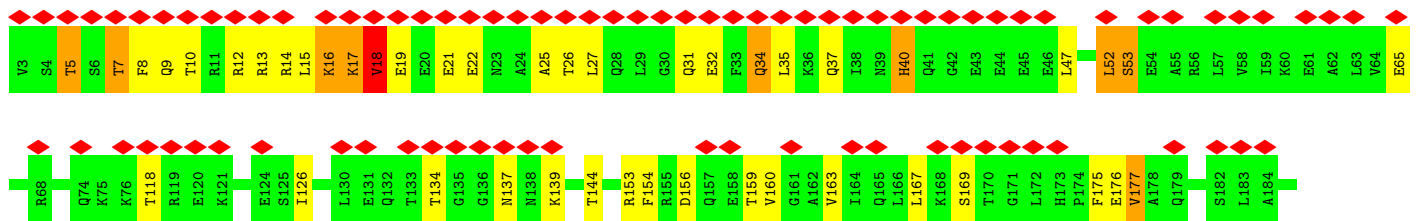
● Molecule 7: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC1, RPB5

Chain E:

43%

74%   24%

D2 Q3 E4 N5 E6 R7 N8 I9 S10 R14 A15 V19 K20 E21 M22 V23 K24 D25 T31 Q32 E33 E34 V35 E36 L37 P38 L39 E40 D41 K45 D48 S49 M50 G51 R52 M57 P64 T65 E66 E67 S68 I69 S70 K71 F72 P73 D74 D84 E85 P86 S87 V88

G89 V90 K91 T92 M93 K94 T95 I98 E102 K103 N104 Q113 N114 N115 I116 T117 P118 S119 K122 L123 V124 P125 P128 P129 A130 T131 I132 E133 T134 F135 N136 E137 A138 A139 L140 N143 I144 T145 H146 H147 E148 L149 V150 I154 R155 L156 S157 S158 D159 E160 K161 R162

K166 R167 S173 P176 R177 I178 Q179 R180 A181 D182 P183 V184 A185 L186 Y187 L188 G189 L190 K191 R192 G193 E194 V195 V196 K197 K201 S202 E203 T204 S205 A209 S210 Y211 R212 M215

● Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC2, RPB6

Chain F:

27%

60%   37%

K72 A73 I74 P75 K76 D77 Q78 R79 T82 P83 Y84 M85 T86 K87 Y88 E89 R90 A91 R92 I93 T96 R97 A98 L99 Q100 I101 S102 M103 N104 A105 P106 V109 D110 L111 E112 G113 E114 T115 D116 P117 L118 R119 M122 E127 I130 P131 L132 V133 I134 R135 R136 Y137 D140

G141 S142 F143 E144 V148 V153 D154 L155

● Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7

Chain G:

54%

67%   27%   5%

M1 F2 K5 D6 L7 S8 L9 N10 I11 T12 L13 H14 P15 S16 F17 F18 G19 P20 Q24 Y25 L26 K29 L30 L31 E32 E33 V34 E35 G36 S37 C38 T39 G40 K41 F42 G43 Y44 I45 L46 D50 Y51 D52 N53 I54 D55 I56 Q57 R58 G59 R60 I61 L62 P63 T64 D65 G66

S67 A68 E69 F70 N71 V72 K73 Y74 V78 F79 K80 P81 F82 K83 G84 E85 V86 V87 D88 S93 Q96 H97 G98 F99 E100 V103 G104 P105 M106 V110 T111 K112 H113 L114 M115 P116 Q117 D118 L119 T120 A123 G124 S125 N126 Q131 S132 S133 E134 D135 V136 I137 T138 I139

R142 I143 R144 V145 K146 I147 E148 G149 C150 I151 S152 Q153 V154 S155 S156 I157 H158 A159 I160 G161 S162 I163 K164 E165 D166 Y167 L168 G169 A170 I171

● Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC3, RPB8
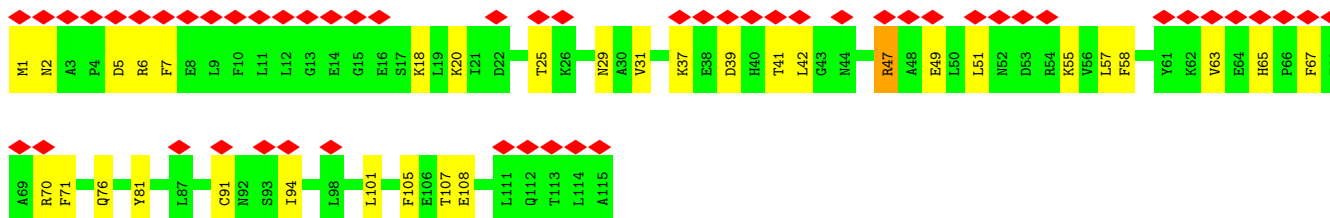
Chain H:

35%

57%   29%   5%   9%
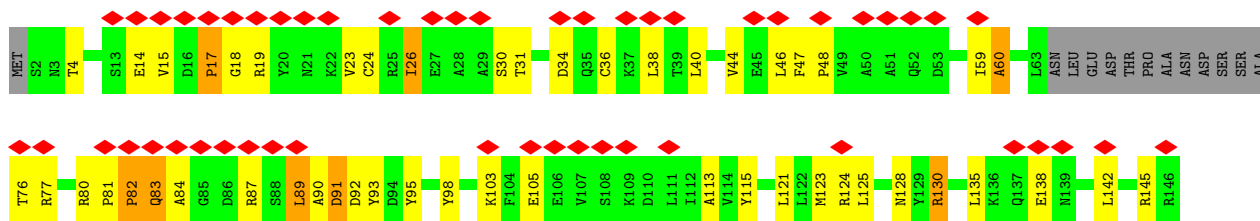
• Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

Chain I:
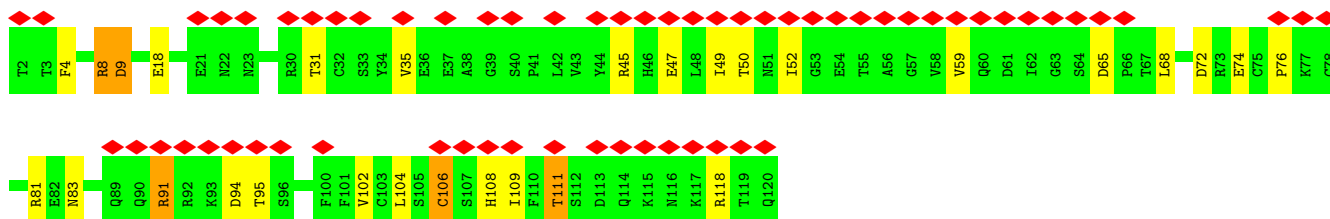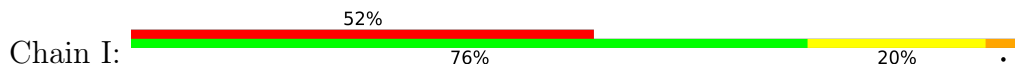
• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC5, RPB10

Chain J:

• Molecule 13: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11
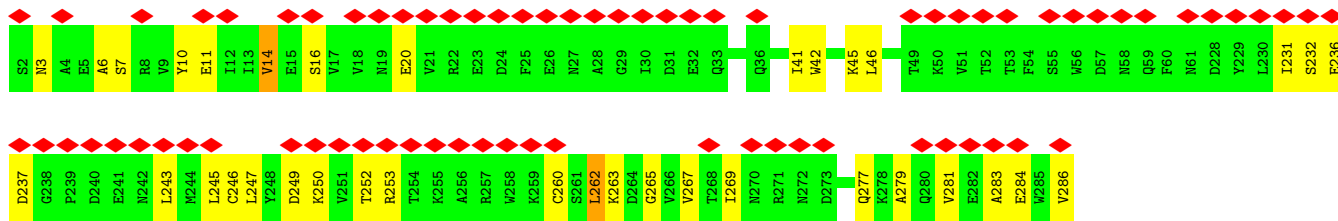
Chain K:

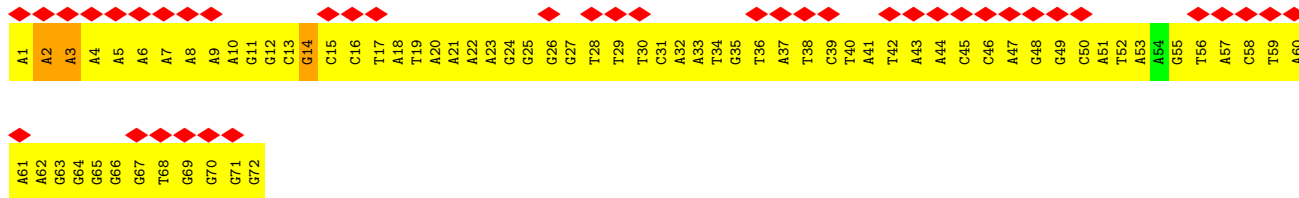• Molecule 14: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC4, RPB12

Chain L:

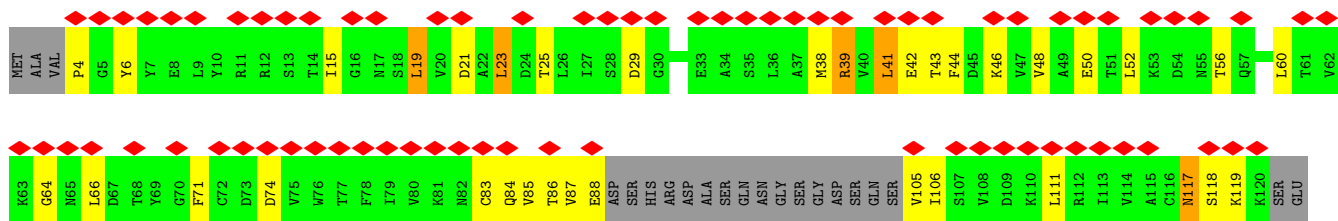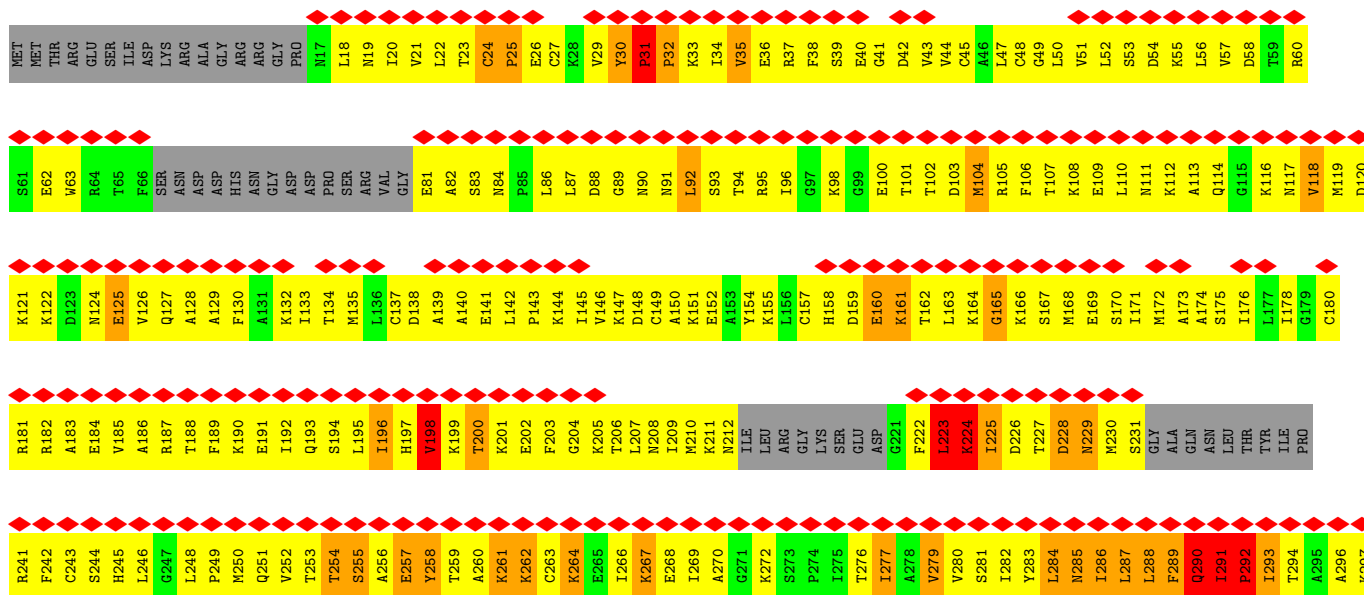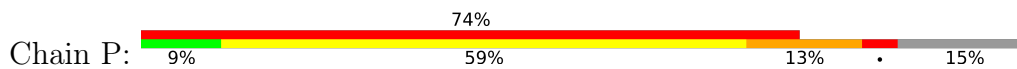• Molecule 15: RNA POLYMERASE II PRE-INITIATION COMPLEX, TOA1

Chain M:

S2 N3 A4 E5 A6 S7 R8 V9 Y10 E11 I12 I13 V14 E15 S16 V17 V18 N19 E20 V21 R22 E23 D24 F25 E26 N27 A28 G29 I30 D31 E32 Q33 Q36 I41 W42 K45 L46 T49 K50 V51 T52 T53 F54 S55 W56 D57 N58 Q59 F60 N61 D228 Y229 L230 I231 S232 E236

D237 G238 F239 D240 E241 N242 L243 M244 L245 C246 L247 Y248 D249 K250 V251 T252 R253 T254 K255 A256 R257 W258 K259 C260 S261 L262 K263 D264 G265 V266 V267 T268 I269 N270 R271 N272 D273 Q277 K278 A279 Q280 V281 E282 A283 E284 W285 V286

• Molecule 16: NON-TEMPLATE STRAND DNA

Chain N:

56%

94%

A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 G11 G12 C13 G14 C15 C16 T17 A18 T19 A20 A21 A22 A23 G24 G25 G26 G27 T28 T29 T30 C31 A32 A33 T34 G35 T36 A37 T38 C39 T40 A41 T42 A43 A44 C45 C46 A47 G48 G49 C50 A51 T52 A53 A54 G55 T56 A57 C58 T59 A60

A61 A62 G63 G64 G65 G66 G67 T68 G69 G70 G71 G72

• Molecule 17: TRANSCRIPTION INITIATION FACTOR IIA SUBUNIT 2, TOA2

Chain O:

61%

53% 25% 17%

MET ALA VAL P4 G5 Y6 Y7 E8 L9 Y10 R11 R12 S13 T14 I15 G16 N17 S18 L19 V20 D21 A22 L23 D24 T25 L26 I27 S28 D29 G30 E33 A34 S35 L36 A37 M38 R39 V40 L41 E42 T43 F44 D45 K46 V47 V48 A49 E50 T51 L52 K53 D54 N55 T56 Q57 L60 T61 V62

K63 G64 N65 L66 D67 T68 Y69 G70 F71 C72 D73 D74 V75 W76 T77 F78 I79 V80 K81 N82 C83 Q84 V85 T86 V87 E88 ASP SER HIS ARG ASP ALA SER GLN ASN GLY SER GLY ASP SER GLN SER V105 I106 S107 V108 D109 K110 L111 R112 I113 V114 A115 C116 N117 S118 K119 K120 SER GLU

• Molecule 18: TRANSCRIPTION INITIATION FACTOR IIB, SUA7

Chain P:

74%

9% 59% 13% 15%

MET MET THR ARG GLU SER ILE ASP LYS ARG ALA GLY ARG ARG GLY PRO N17 L18 N19 I20 V21 L22 T23 C24 P25 E26 C27 K28 V29 Y30 P31 P32 K33 I34 V35 E36 R37 F38 S39 E40 G41 D42 V43 V44 C45 A46 L47 C48 G49 L50 V51 L52 S53 D54 K55 L56 V57 D58 T59 R60

S61 E62 W63 R64 T65 F66 SER ASN ASP ASP HIS ASN GLY ASP ASP PRO SER ARG VAL GLY E81 A82 S83 N84 P85 L86 L87 D88 G89 N90 N91 L92 S93 T94 R95 I96 G97 K98 G99 E100 T101 T102 D103 M104 R105 F106 T107 K108 E109 L110 N111 K112 A113 Q114 G115 K116 N117 V118 M119 D120

K121 K122 D123 N124 E125 V126 Q127 A128 A129 F130 A131 K132 I133 T134 M135 L136 C137 D138 A139 A140 E141 L142 P143 K144 I145 V146 N147 D148 C149 A150 K151 E152 A153 Y154 K155 L156 C157 H158 D159 E160 K161 T162 L163 K164 G165 K166 S167 M168 E169 S170 I171 M172 A173 A174 S175 I176 L177 I178 G179 C180

R181 R182 A183 E184 V185 A186 R187 T188 F189 K190 E191 I192 Q193 S194 L195 I196 H197 V198 K199 T200 K201 E202 F203 G204 K205 T206 L207 N208 I209 M210 K211 N212 ILE LEU ARG GLY LYS SER GLU ASP G221 F222 L223 K224 I225 D226 T227 D228 N229 M230 S231 GLY ALA GLN ASN LEU THR TYR ILE PRO

R241 F242 C243 S244 H245 L246 G247 L248 P249 M250 Q251 V252 T253 T254 S255 A256 E257 Y258 T259 A260 K261 K262 C263 K264 E265 I266 K267 E268 I269 A270 G271 K272 S273 P274 I275 T276 I277 A278 V279 V280 S281 I282 Y283 L284 N285 I286 L287 L288 F289 Q290 I291 P292 I293 T294 A295 A296 K297 V298 G299 Q300

T301 L302 Q303 V304 T305 E306 G307 T308 I309 K310 S311 G312 Y313 K314 I315 L316 Y317 E318 H319 R320 D321 K322 L323 V324 D325 P326 Q327 L328 I329 A330 N331 G332 V333 V334 S335 L336 D337 N338 L339 P340 GLY VAL GLU LYS LYS

• Molecule 19: TATA-BOX-BINDING PROTEIN, TBP

Chain Q:

53%
59%　　　　　36%

S61 G62 I63 V64 P65 T66 L67 Q68 T75 L76 G77 C78 R79 L80 D81 L82 K83 T84 V85 A86 L87 H88 A89 R90 E93 Y94 N95 P96 K97 R98 F99 A100 I103 I106 R107 E108 P109 K110 T111 T112 A113 L114 I115 F116 A117 S118 G119 K120 M121 V122 V123 T124 G125 A126 K127 S128

E129 D130 D131 S132 K133 L134 A135 S136 R137 K138 Y139 A140 R141 I142 I143 Q144 K145 I146 G147 F148 A149 A150 K151 F152 T153 D154 F155 K156 I157 Q158 N159 G162 S163 C164 D165 V166 K167 F168 P169 I170 R171 L172 E173 G174 L175 A176 F177 S178 H179 G180 T181 F182 S183 E186 P187 E188 L189

L193 I194 Y195 R196 M197 V198 K199 I202 V203 L204 L205 V208 S209 G210 K211 I212 V213 L214 T215 G216 A217 K218 Q219 R220 E221 Y224 Q225 E228 A229 I230 Y231 P232 V233 L234 S235 E236 F237 R238 K239 M240

• Molecule 20: TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT ALPHA, TFA1

Chain R:

38%
14%　　　46%　　　11%　　26%

MET ASP ARG PRO ILE ASP ASP ILE VAL LYS ASN L12 L13 K14 F15 V16 V17 R18 G19 F20 Y21 F25 V26 L27 V28 N29 D30 A31 I32 L33 F34 H35 S36 V37 L38 A39 E40 D41 D42 L43 K44 Q45 L46 L47 S48 I49 N50 K51 T52 E53 L54 G55 P56 L57 I58 A59 R60 L61 R62

S63 D64 R65 L66 I67 S68 I69 H70 K71 Q72 R73 GLU TYR PRO PRO ASN SER LYS SER V82 E83 R84 V85 Y86 Y87 Y88 V89 K90 Y91 P92 H93 A94 I95 ASP ALA ILE LYS TRP LYS VAL HIS GLN VAL GLN ARG LEU LYS ASP ASP LEU ASP LYS ASN SER GLU P119 N120 G121 Y122

M123 C124 P125 I126 C127 L128 T129 K130 Y131 T132 Q133 L134 E135 A136 V137 Q138 L139 L140 N141 F142 D143 R144 T145 E146 F147 L148 C149 S150 L151 C152 D153 E154 P155 L156 V157 E158 D159 D160

• Molecule 21: TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT BETA, TFA2

Chain S:

44%
14%　　　49%　　　28%　　7%

S125 S126 K127 L128 L129 W130 A131 T132 E133 Y134 I135 Q136 K137 K138 G139 K140 P141 V142 L143 V144 N145 E146 L147 L148 D149 Y150 L151 S152 M153 K154 K155 D156 D157 K158 V159 I160 E161 L162 L163 K164 K165 L166 D167 R168 I169 E170 F171 D172 PRO LYS K175 G176 T177 F178 K179 Y180 L181 S182 T183 Y184

D185 V186 H187 S188 P189 S190 E191 L192 L193 K194 L195 L196 R197 S198 Q199 V200 T201 P202 K203 G204 I205 S206 C207 K208 D209 L210 K211 D212 G213 W214 P215 Q216 C217 D218 E219 T220 I221 N222 Q223 L224 E225 E226 D227 S228 K229 I230 L231 V232 L233 R234 K237 D238 K239 T240 P241 R242 Y243 V244 W245

Y246 N247

• Molecule 22: TEMPLATE STRAND DNA

Chain T:

- Molecule 23: RNA POLYMERASE II PRE-INITIATION COMPLEX, TFG1

Chain U:

- Molecule 24: TRANSCRIPTION INITIATION FACTOR IIF SUBUNIT BETA, TFG2

Chain V:
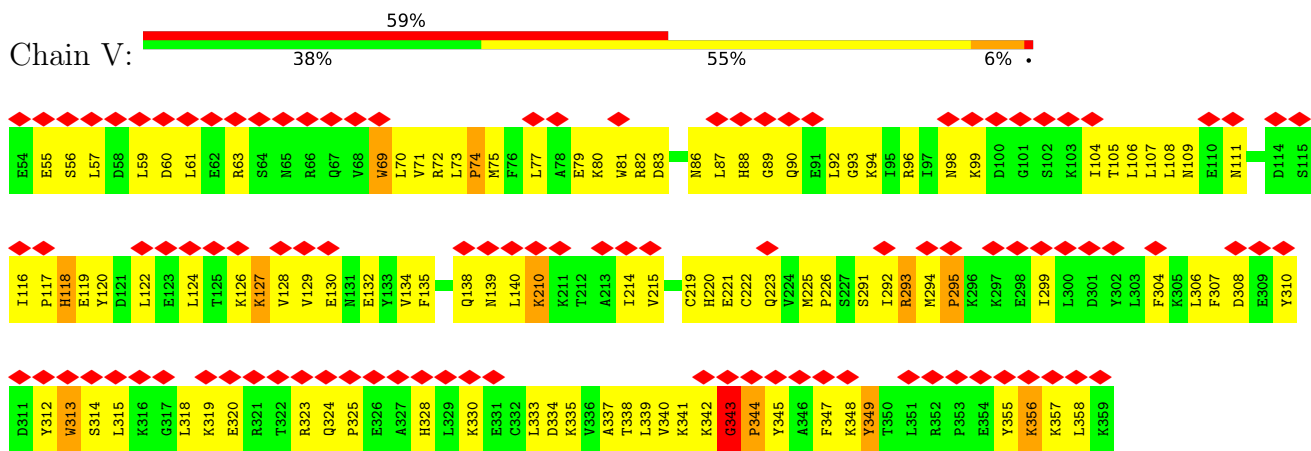
- Molecule 25: RNA POLYMERASE II TRANSCRIPTION FACTOR B SUBUNIT 2, TFB2

Chain W:
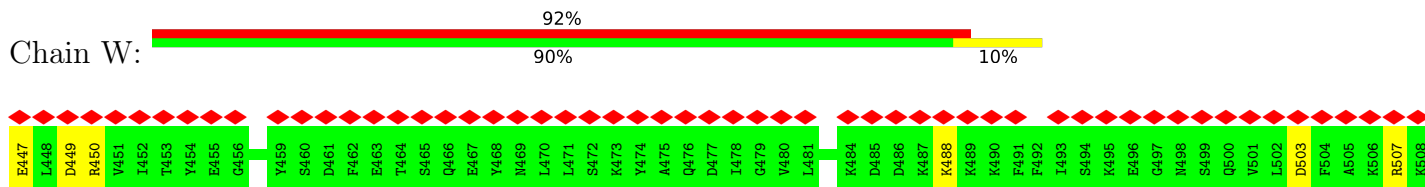
- Molecule 26: RNA POLYMERASE II TRANSCRIPTION FACTOR B SUBUNIT 5, TFB5

Chain X:
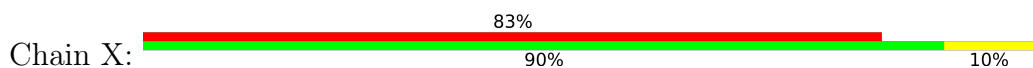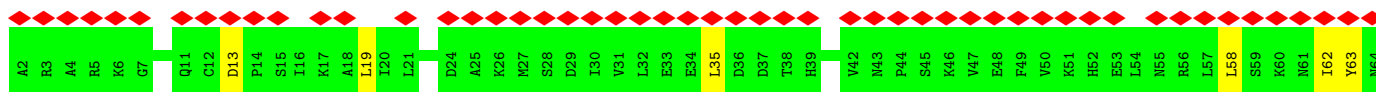
• Molecule 27: DNA REPAIR HELICASE RAD3

Chain Y:

27%

33%   38%   26%
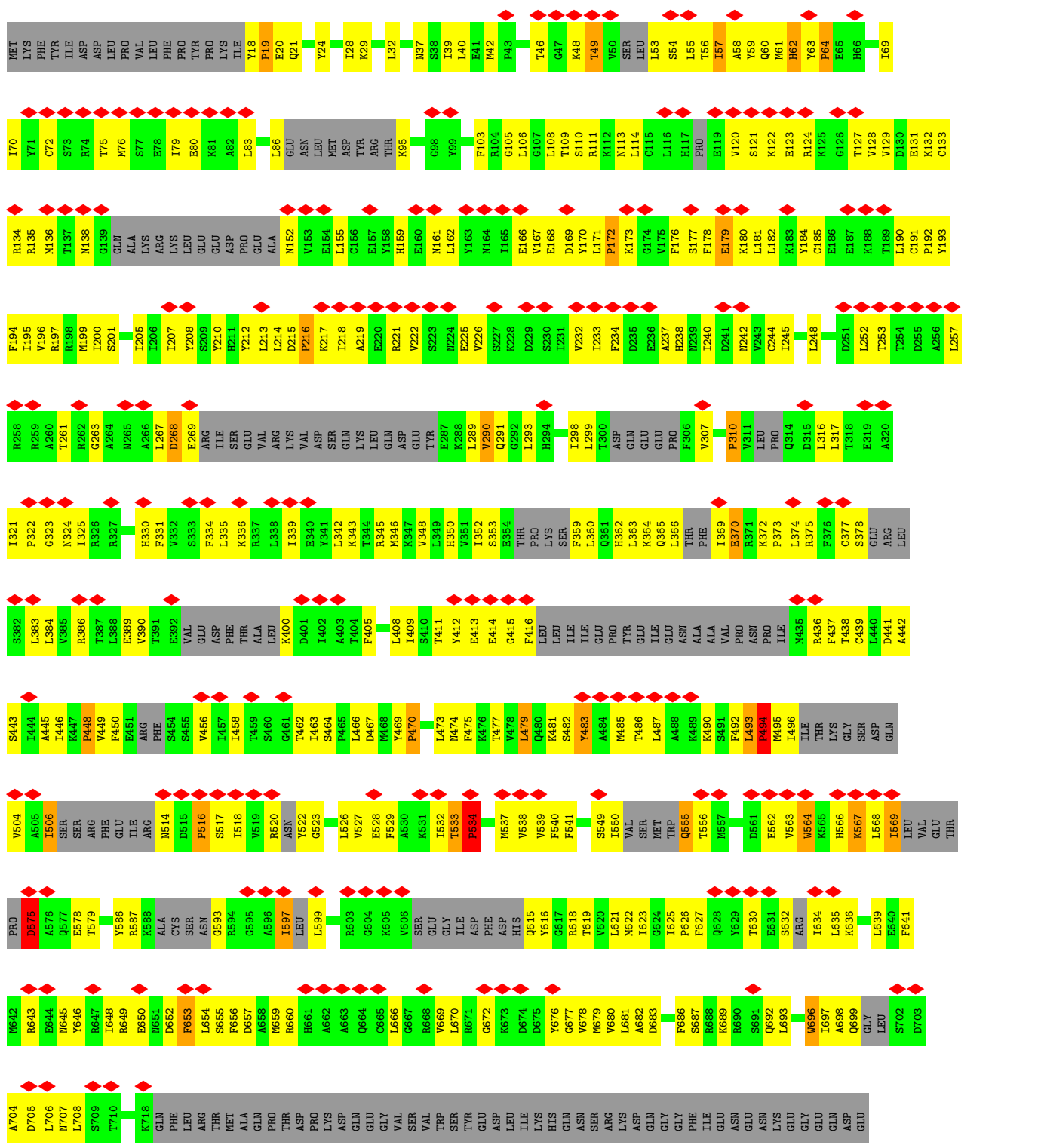
ASP GLU ASP GLU ASP ILE GLU MET GLN

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	7578	Depositor
Resolution determination method	Not provided	
CTF correction method	SPARX	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{Å}^2$)	40	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	22500	Depositor
Image detector	GATAN K2 (4k x 4k)	Depositor
Maximum map value	9.151	Depositor
Minimum map value	-0.003	Depositor
Average map value	0.142	Depositor
Map value standard deviation	0.652	Depositor
Recommended contour level	3.99	Depositor
Map size (Å)	399.76, 399.76, 399.76	wwPDB
Map dimensions	152, 152, 152	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	2.63, 2.63, 2.63	Depositor

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#|Z| > 5$	RMSZ	$\#|Z| > 5$
1	1	4.34	6/3867 (0.2%)	2.05	41/5221 (0.8%)
2	2	1.31	4/571 (0.7%)	1.64	7/765 (0.9%)
3	A	0.57	4/11374 (0.0%)	0.81	11/15384 (0.1%)
4	B	0.49	1/9316 (0.0%)	0.74	4/12564 (0.0%)
5	C	0.49	0/2133	0.78	2/2891 (0.1%)
6	D	0.51	0/1444	0.83	2/1935 (0.1%)
7	E	0.48	0/1788	0.72	0/2406
8	F	0.62	0/691	0.81	0/933
9	G	0.52	0/1368	0.81	0/1844
10	H	0.51	0/1086	0.80	0/1470
11	I	0.47	0/989	0.78	0/1331
12	J	0.54	0/541	0.88	1/727 (0.1%)
13	K	0.47	0/938	0.71	0/1267
14	L	0.54	0/365	0.95	0/485
15	M	0.40	0/970	0.57	0/1310
16	N	1.27	20/1692 (1.2%)	1.41	34/2613 (1.3%)
17	O	0.42	0/800	0.63	0/1080
18	P	4.21	7/2297 (0.3%)	2.23	41/3090 (1.3%)
19	Q	0.41	0/1443	0.62	0/1942
20	R	9.51	2/978 (0.2%)	3.28	11/1321 (0.8%)
21	S	5.35	7/1017 (0.7%)	3.30	36/1356 (2.7%)
22	T	1.38	45/1618 (2.8%)	1.50	49/2491 (2.0%)
23	U	0.93	4/1273 (0.3%)	0.92	9/1710 (0.5%)
24	V	1.09	3/1481 (0.2%)	0.88	10/1988 (0.5%)
25	W	0.67	0/527	0.68	0/704
26	X	0.60	0/504	0.69	0/679
27	Y	0.88	11/4735 (0.2%)	0.84	19/6369 (0.3%)
All	All	2.13	114/55806 (0.2%)	1.23	277/75876 (0.4%)

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	1	1
2	2	0	2
3	A	0	2
4	B	0	1
18	P	2	0
20	R	1	0
21	S	0	1
27	Y	0	2
All	All	4	9

The worst 5 of 114 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
20	R	155	PRO	N-CD	213.23	4.46	1.47
20	R	119	PRO	N-CD	204.92	4.34	1.47
1	1	753	PRO	N-CD	203.03	4.32	1.47
18	P	292	PRO	N-CD	198.01	4.25	1.47
1	1	746	PRO	N-CD	168.34	3.83	1.47

The worst 5 of 277 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
20	R	155	PRO	CA-N-CD	-66.71	18.11	111.50
20	R	119	PRO	N-CD-CG	-58.68	15.18	103.20
21	S	189	PRO	CA-N-CD	-53.80	36.18	111.50
1	1	753	PRO	N-CD-CG	-48.67	30.19	103.20
18	P	292	PRO	N-CD-CG	-48.02	31.17	103.20

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	1	690	ILE	CB
18	P	269	ILE	CB
18	P	277	ILE	CB
20	R	126	ILE	CB

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1	661	SER	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	2	269	PHE	Peptide,Sidechain
3	A	234	MET	Mainchain
3	A	95	PHE	Mainchain
4	B	43	LEU	Mainchain

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	1	3797	0	3727	916	0
2	2	666	0	546	251	0
3	A	11174	0	11217	712	0
4	B	9140	0	9107	421	0
5	C	2095	0	2051	43	0
6	D	1434	0	1460	69	0
7	E	1752	0	1776	30	0
8	F	679	0	699	94	0
9	G	1340	0	1353	174	0
10	H	1068	0	1040	21	0
11	I	971	0	927	16	0
12	J	532	0	542	14	0
13	K	920	0	929	19	0
14	L	363	0	386	24	0
15	M	956	0	916	32	0
16	N	1500	0	805	235	0
17	O	792	0	806	49	0
18	P	2269	0	2346	1242	0
19	Q	1416	0	1491	154	0
20	R	960	0	969	345	0
21	S	1005	0	1024	580	0
22	T	1452	0	822	273	0
23	U	1248	0	1229	236	0
24	V	1455	0	1465	465	0
25	W	518	0	514	27	0
26	X	499	0	525	12	0
27	Y	4660	691	4743	807	0
28	2	1	0	0	0	0
28	A	1	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
29	2	1	0	0	0	0
29	A	2	0	0	0	0
29	B	1	0	0	0	0
29	C	1	0	0	0	0
29	I	2	0	0	0	0
29	J	1	0	0	0	0
29	L	1	0	0	0	0
All	All	54672	691	53415	6017	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 56.

The worst 5 of 6017 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)
22:T:134:DT:C6	24:V:323:ARG:HD2	1.24	1.69
18:P:279:VAL:HG23	18:P:302:LEU:CG	1.19	1.66
20:R:34:PHE:CE2	20:R:134:LEU:HD11	1.25	1.65
18:P:137:CYS:SG	18:P:142:LEU:HB3	1.34	1.64
27:Y:132:LYS:HD3	27:Y:155:LEU:CD1	1.24	1.64

There are no symmetry-related clashes.

5.3   Torsion angles   ⓘ

5.3.1   Protein backbone   ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	1	467/492 (95%)	436 (93%)	26 (6%)	5 (1%)	14	52
2	2	68/174 (39%)	51 (75%)	10 (15%)	7 (10%)	0	8
3	A	1414/1733 (82%)	1250 (88%)	113 (8%)	51 (4%)	3	25
4	B	1140/1224 (93%)	1018 (89%)	86 (8%)	36 (3%)	4	26

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	C	264/266 (99%)	242 (92%)	20 (8%)	2 (1%)	19	60
6	D	174/178 (98%)	148 (85%)	18 (10%)	8 (5%)	2	21
7	E	212/214 (99%)	195 (92%)	13 (6%)	4 (2%)	8	38
8	F	82/84 (98%)	75 (92%)	7 (8%)	0	100	100
9	G	169/171 (99%)	158 (94%)	8 (5%)	3 (2%)	8	40
10	H	129/146 (88%)	106 (82%)	14 (11%)	9 (7%)	1	14
11	I	117/119 (98%)	98 (84%)	16 (14%)	3 (3%)	5	31
12	J	63/65 (97%)	51 (81%)	9 (14%)	3 (5%)	2	20
13	K	113/115 (98%)	109 (96%)	4 (4%)	0	100	100
14	L	44/46 (96%)	27 (61%)	9 (20%)	8 (18%)	0	2
15	M	110/116 (95%)	103 (94%)	7 (6%)	0	100	100
17	O	97/122 (80%)	93 (96%)	4 (4%)	0	100	100
18	P	287/345 (83%)	235 (82%)	33 (12%)	19 (7%)	1	15
19	Q	178/180 (99%)	170 (96%)	5 (3%)	3 (2%)	9	42
20	R	112/160 (70%)	100 (89%)	10 (9%)	2 (2%)	8	40
21	S	105/123 (85%)	94 (90%)	7 (7%)	4 (4%)	3	24
23	U	142/150 (95%)	131 (92%)	6 (4%)	5 (4%)	3	25
24	V	166/174 (95%)	145 (87%)	16 (10%)	5 (3%)	4	28
25	W	60/62 (97%)	60 (100%)	0	0	100	100
26	X	61/63 (97%)	58 (95%)	3 (5%)	0	100	100
27	Y	558/778 (72%)	515 (92%)	35 (6%)	8 (1%)	11	46
All	All	6332/7300 (87%)	5668 (90%)	479 (8%)	185 (3%)	7	29

5 of 185 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	447	GLN
1	1	458	SER
2	2	275	LYS
2	2	292	PRO
3	A	74	MET

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	1	415/431 (96%)	370 (89%)	45 (11%)	6	24
2	2	62/151 (41%)	55 (89%)	7 (11%)	6	22
3	A	1240/1520 (82%)	1066 (86%)	174 (14%)	3	17
4	B	985/1061 (93%)	868 (88%)	117 (12%)	5	20
5	C	234/234 (100%)	206 (88%)	28 (12%)	5	20
6	D	160/160 (100%)	129 (81%)	31 (19%)	1	8
7	E	196/196 (100%)	175 (89%)	21 (11%)	6	24
8	F	74/74 (100%)	67 (90%)	7 (10%)	8	28
9	G	152/152 (100%)	135 (89%)	17 (11%)	6	22
10	H	117/128 (91%)	103 (88%)	14 (12%)	5	20
11	I	113/113 (100%)	106 (94%)	7 (6%)	18	43
12	J	60/60 (100%)	49 (82%)	11 (18%)	1	10
13	K	99/99 (100%)	87 (88%)	12 (12%)	5	20
14	L	40/40 (100%)	27 (68%)	13 (32%)	0	2
15	M	107/107 (100%)	104 (97%)	3 (3%)	43	65
17	O	91/108 (84%)	83 (91%)	8 (9%)	10	31
18	P	256/299 (86%)	215 (84%)	41 (16%)	2	13
19	Q	152/152 (100%)	143 (94%)	9 (6%)	19	45
20	R	108/149 (72%)	88 (82%)	20 (18%)	1	9
21	S	116/118 (98%)	69 (60%)	47 (40%)	0	0
23	U	136/136 (100%)	133 (98%)	3 (2%)	52	71
24	V	163/163 (100%)	158 (97%)	5 (3%)	40	62
25	W	57/57 (100%)	57 (100%)	0	100	100
26	X	57/57 (100%)	57 (100%)	0	100	100
27	Y	521/707 (74%)	504 (97%)	17 (3%)	38	61
All	All	5711/6472 (88%)	5054 (88%)	657 (12%)	9	21

5 of 657 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	I	35	VAL
20	R	133	GLN
12	J	42	LYS
11	I	31	THR
18	P	255	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 109 such sidechains are listed below:

Mol	Chain	Res	Type
10	H	35	GLN
18	P	84	ASN
27	Y	60	GLN
11	I	83	ASN
15	M	59	GLN

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 ⓘ

Of 11 ligands modelled in this entry, 11 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7　Other polymers　ⓘ

There are no such residues in this entry.

5.8　Polymer linkage issues　ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
21	S	7
27	Y	7
24	V	3
23	U	3
15	M	2
3	A	2
6	D	1
1	1	1
4	B	1
2	2	1

The worst 5 of 28 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	V	227:SER	C	291:SER	N	84.31
1	M	61:ASN	C	228:ASP	N	43.11
1	D	76:LYS	C	118:THR	N	35.68
1	S	184:TYR	C	185:ASP	N	14.79
1	U	313:ALA	C	323:GLY	N	13.13

6   Map visualisation   ⓘ

This section contains visualisations of the EMDB entry EMD-3114. 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 |
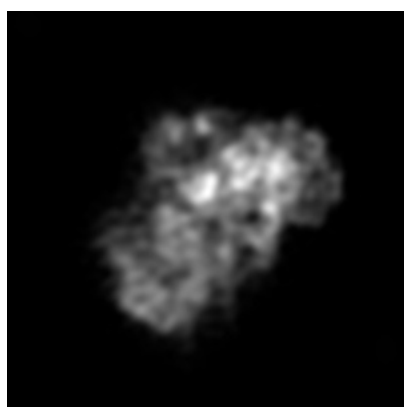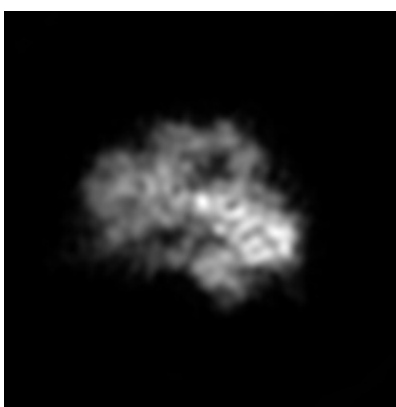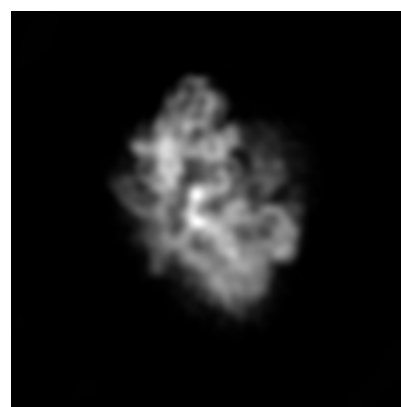
The images above show the map projected in three orthogonal directions.

6.2   Central slices   ⓘ
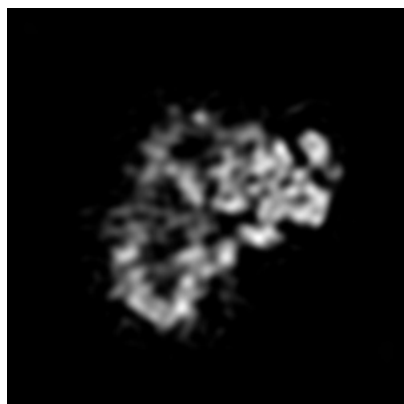
6.2.1   Primary map
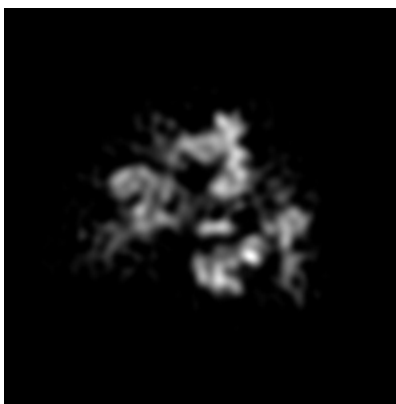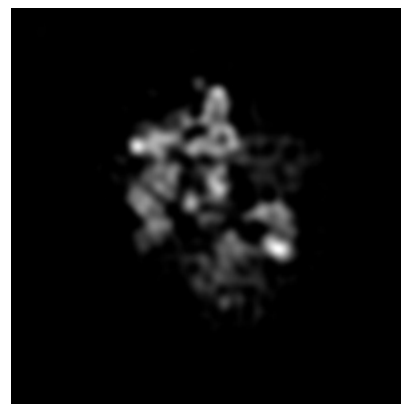
| X Index: 76 | Y Index: 76 | Z Index: 76 |

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

6.3 Largest variance slices ⓘ
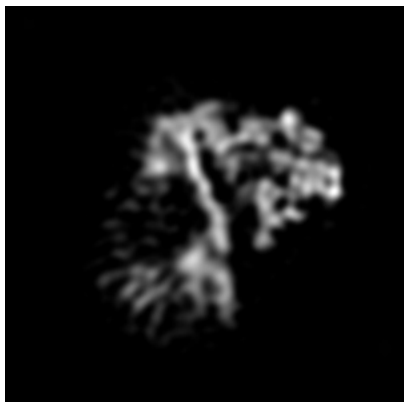
6.3.1 Primary map
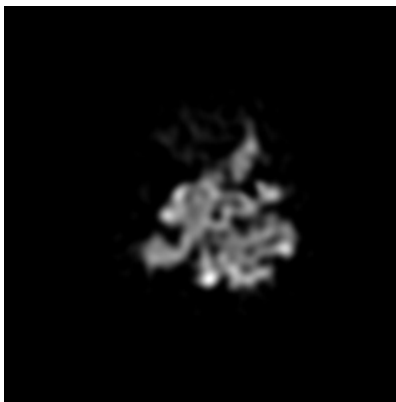
| X Index: 69 | Y Index: 99 | Z Index: 88 |

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

6.4 Orthogonal surface views ⓘ
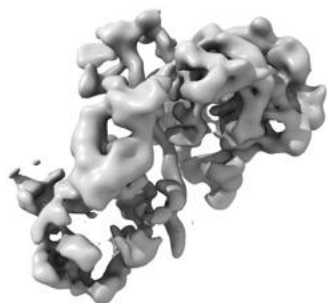
6.4.1 Primary map

| X | Y | Z |
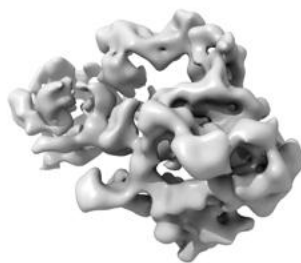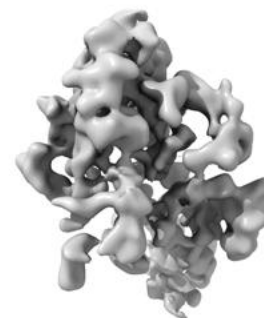
The images above show the 3D surface view of the map at the recommended contour level 3.99. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

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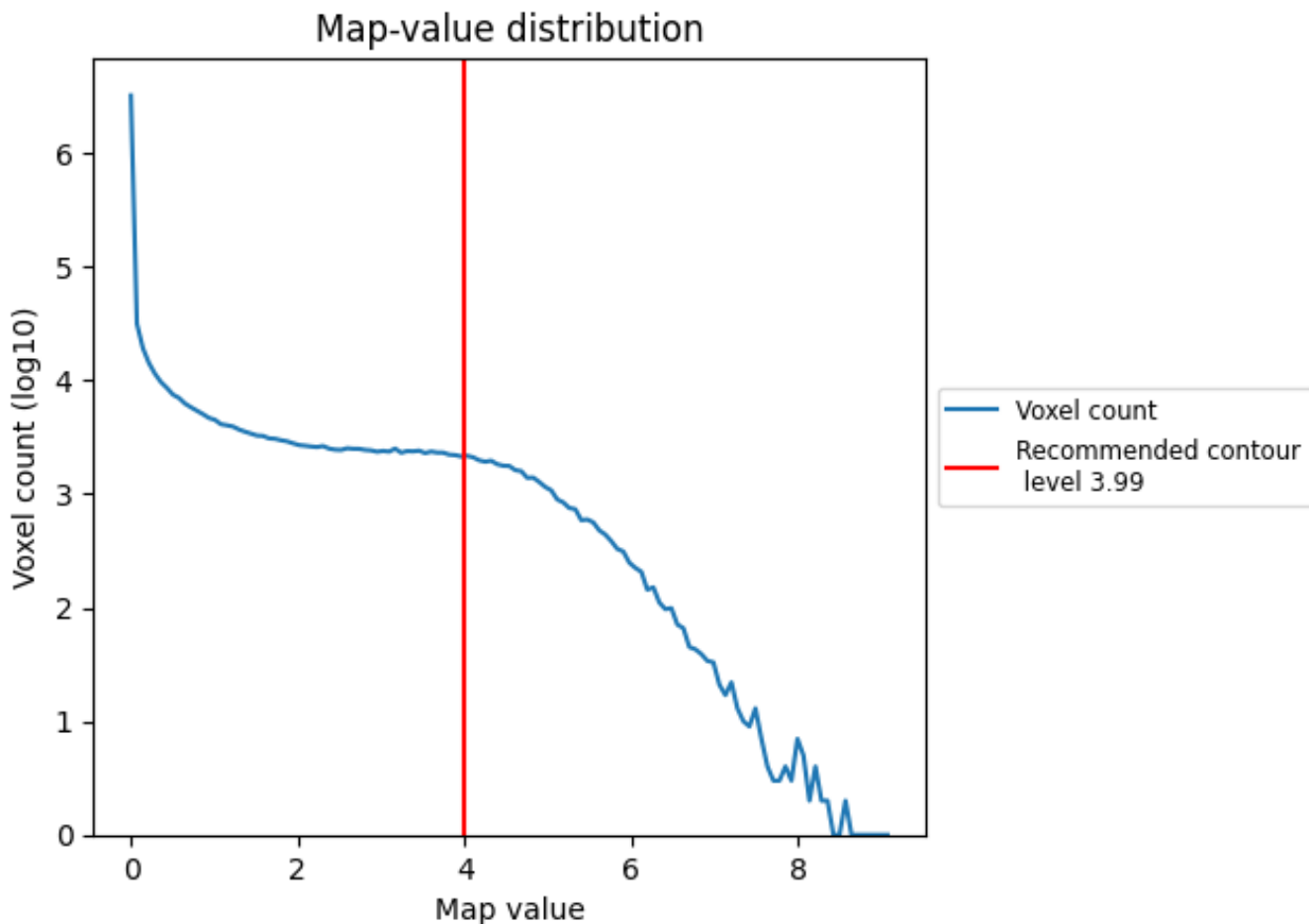
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2   Volume estimate ⓘ
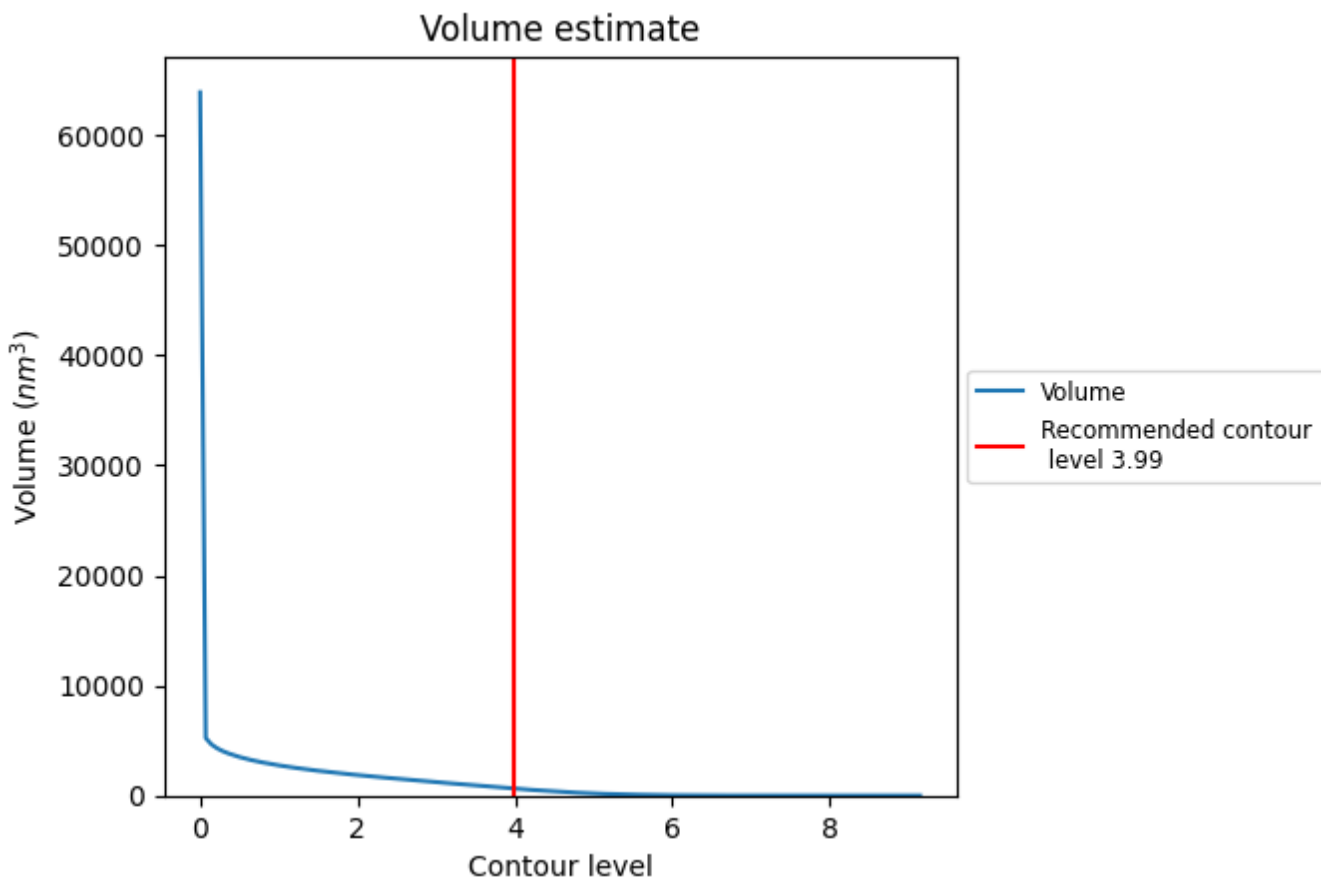
The volume at the recommended contour level is 637 nm^3; this corresponds to an approximate mass of 575 kDa.

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)
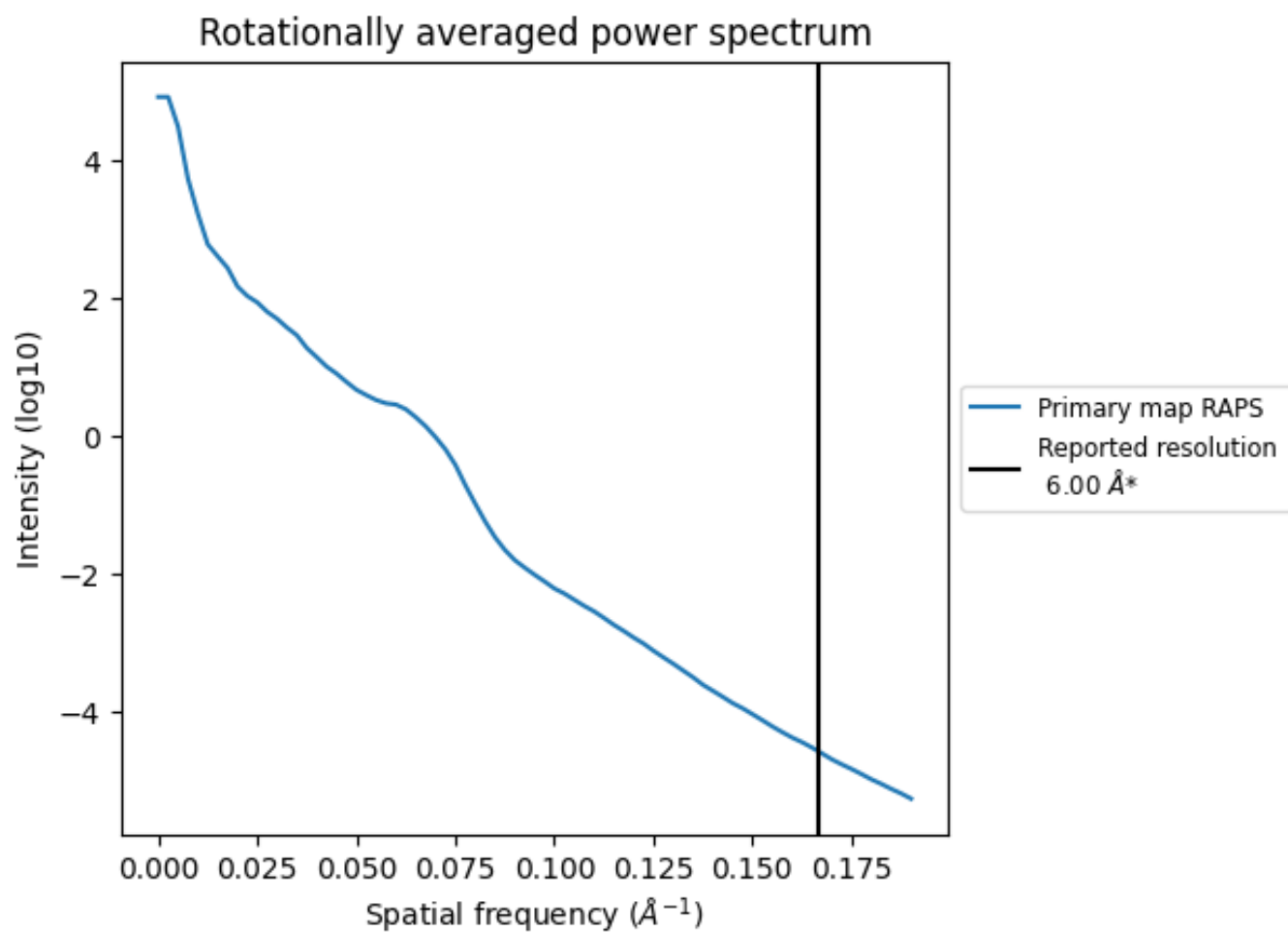
*Reported resolution corresponds to spatial frequency of 0.167 Å^{-1}

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-3114 and PDB model 5FMF. Per-residue inclusion information can be found in section 3 on page 9.

9.1　Map-model overlay ⓘ

| X | Y | Z |
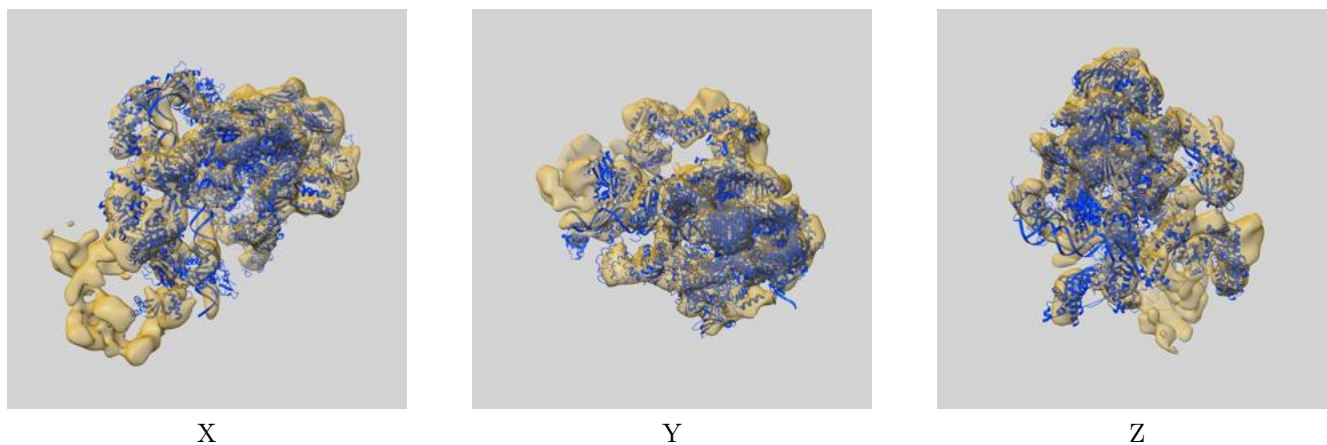
The images above show the 3D surface view of the map at the recommended contour level 3.99 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2   Q-score mapped to coordinate model ⓘ
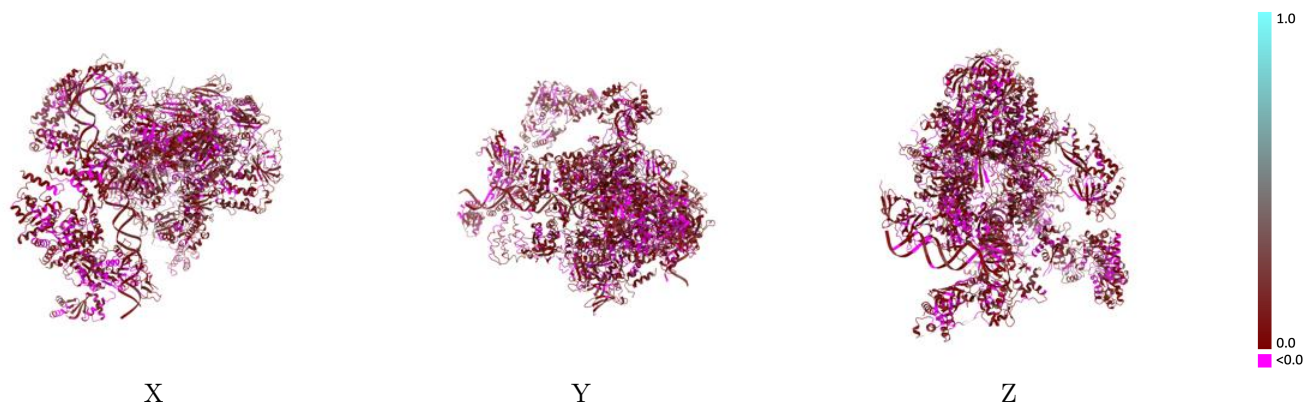
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 ⓘ
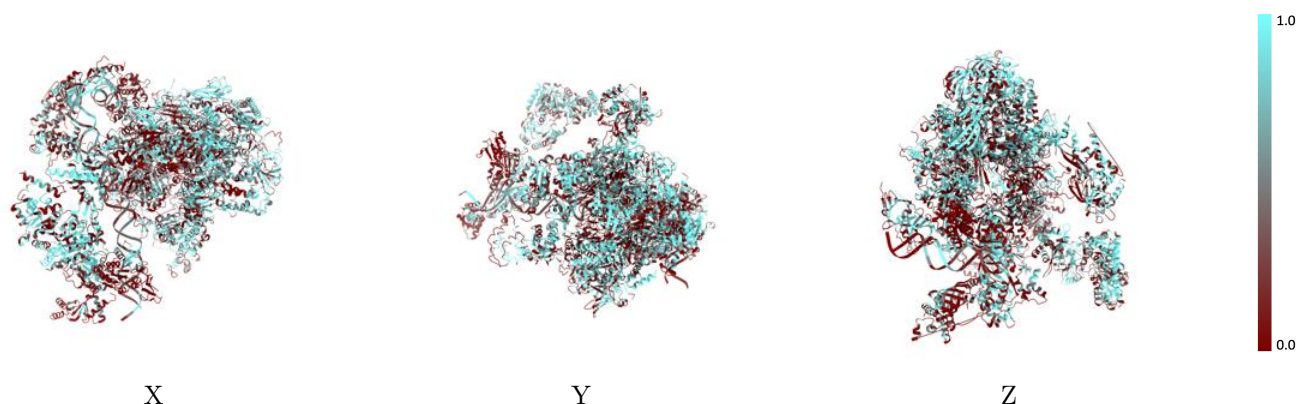
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (3.99).

9.4   Atom inclusion ⓘ
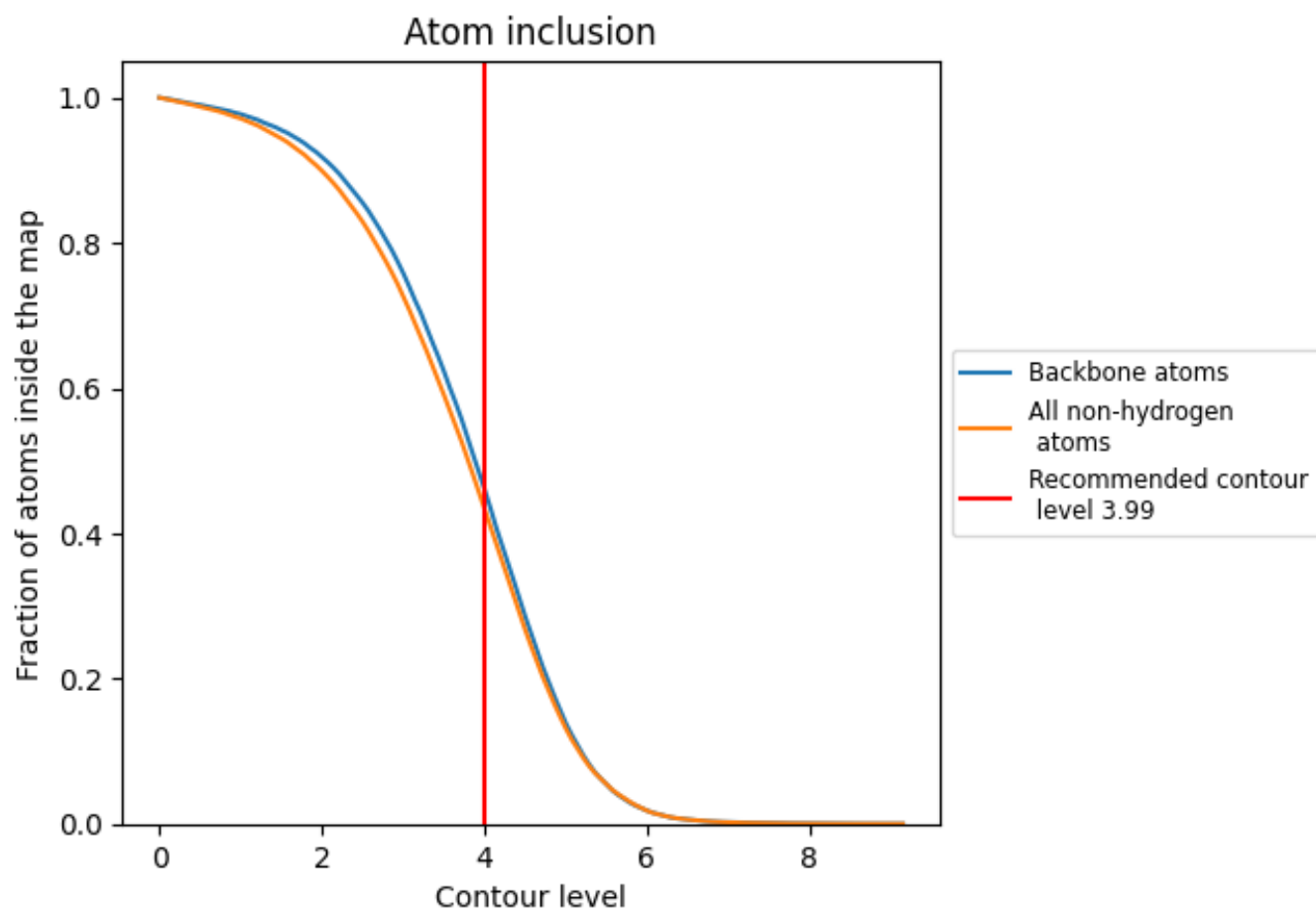
At the recommended contour level, 46% of all backbone atoms, 44% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (3.99) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.4368	0.0620
1	0.2962	0.0410
2	0.2866	0.0410
A	0.5002	0.0770
B	0.4001	0.0490
C	0.6132	0.0640
D	0.3741	0.0700
E	0.5307	0.0900
F	0.6778	0.0690
G	0.4183	0.0690
H	0.5760	0.0720
I	0.4408	0.0600
J	0.5841	0.0660
K	0.5138	0.0540
L	0.6553	0.0600
M	0.3348	0.0730
N	0.3933	0.0730
O	0.2420	0.0580
P	0.1213	0.0240
Q	0.4330	0.0640
R	0.4707	0.0730
S	0.5091	0.1020
T	0.5021	0.1110
U	0.3739	0.0320
V	0.3862	0.0590
W	0.0923	0.0590
X	0.1870	0.0240
Y	0.5868	0.0570

1.0

0.0
<0.0