wwPDB EM Validation Summary Report (i)

May 14, 2024 – 12:21 AM JST

PDB ID	:	7EVO
EMDB ID	:	EMD-31334
Title	:	The cryo-EM structure of the human 17S U2 snRNP
Authors	:	Zhang, X.; Zhan, X.; Shi, Y.
Deposited on	:	2021-05-21
Resolution	:	2.50 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (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.dev92
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.50 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.
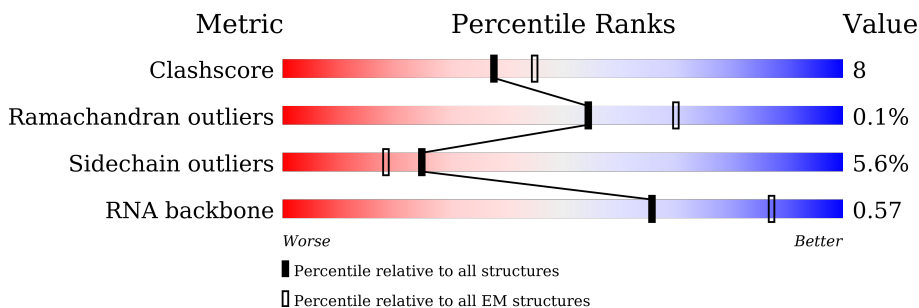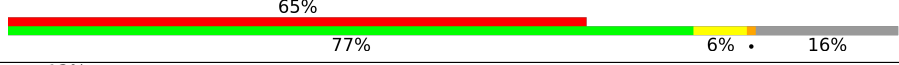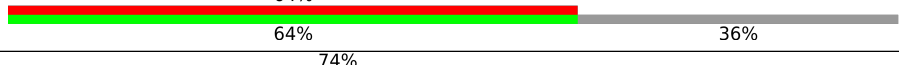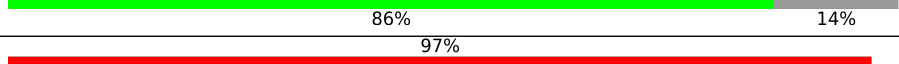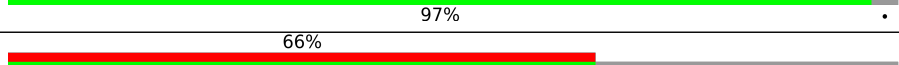

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	H	188	68% / 45% 25% 8% · 21%
2	1	1304	· / 47% 15% · 38%
3	2	895	9% / 17% · 80%
4	3	1217	16% / 73% 23% · ·
5	4	424	38% / 38% 62%
6	5	86	60% 16% 23%
7	6	110	· / 64% 14% 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	A	793	16% / 16% / 84%
9	B	464	22% / 22% / 78%
10	C	501	65% / 77% / 6% · 16%
11	D	755	13% / 22% · 74%
12	E	1052	41% / 41% · 57%
13	F	255	64% / 64% / 36%
14	G	225	74% / 73% / 26%
15	a	118	76% / 76% / 24%
16	b	86	86% / 86% / 14%
17	c	92	86% / 86% / 14%
18	d	76	97% / 97% ·
19	e	126	66% / 66% / 34%
20	f	240	30% / 30% / 70%
21	g	119	69% / 69% / 31%

2   Entity composition ⓘ

There are 23 unique types of molecules in this entry. The entry contains 32104 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 RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	H	148	Total 3142	C 1406	N 551	O 1039	P 146	0	0

- Molecule 2 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	1	815	Total 6487	C 4163	N 1121	O 1164	S 39	0	0

- Molecule 3 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	2	182	Total 1210	C 754	N 229	O 225	S 2	0	0

- Molecule 4 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	3	1180	Total 9247	C 5872	N 1571	O 1759	S 45	0	0

- Molecule 5 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	4	160	Total 664	C 344	N 160	O 160	0	0

- Molecule 6 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	5	66	Total 539	C 342	N 93	O 99	S 5	0	0

- Molecule 7 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	6	85	Total 645	C 396	N 114	O 122	S 13	0	0

- Molecule 8 is a protein called Splicing factor 3A subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	A	123	Total 504	C 258	N 123	O 123	0	0

- Molecule 9 is a protein called Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	B	100	Total 421	C 221	N 100	O 100	0	0

- Molecule 10 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	C	422	Total 2103	C 1174	N 466	O 459	S 4	0	0

- Molecule 11 is a protein called HIV Tat-specific factor 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	D	200	Total 1308	C 788	N 246	O 269	S 5	3	0

- Molecule 12 is a protein called RNA helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	E	455	Total 2153	C 1157	N 487	O 507	S 2	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-19	MET	-	initiating methionine	UNP A0A0C4DG89
E	-18	ALA	-	expression tag	UNP A0A0C4DG89
E	-17	SER	-	expression tag	UNP A0A0C4DG89
E	-16	ASP	-	expression tag	UNP A0A0C4DG89

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-15	TYR	-	expression tag	UNP A0A0C4DG89
E	-14	LYS	-	expression tag	UNP A0A0C4DG89
E	-13	ASP	-	expression tag	UNP A0A0C4DG89
E	-12	ASP	-	expression tag	UNP A0A0C4DG89
E	-11	ASP	-	expression tag	UNP A0A0C4DG89
E	-10	ASP	-	expression tag	UNP A0A0C4DG89
E	-9	LYS	-	expression tag	UNP A0A0C4DG89
E	-8	ALA	-	expression tag	UNP A0A0C4DG89
E	-7	SER	-	expression tag	UNP A0A0C4DG89
E	-6	ASP	-	expression tag	UNP A0A0C4DG89
E	-5	GLU	-	expression tag	UNP A0A0C4DG89
E	-4	VAL	-	expression tag	UNP A0A0C4DG89
E	-3	ASP	-	expression tag	UNP A0A0C4DG89
E	-2	ALA	-	expression tag	UNP A0A0C4DG89
E	-1	GLY	-	expression tag	UNP A0A0C4DG89
E	0	THR	-	expression tag	UNP A0A0C4DG89

- Molecule 13 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	F	162	Total	C	N	O	0	0
			666	342	162	162		

- Molecule 14 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms				AltConf	Trace
14	G	166	Total	C	N	O	0	0
			685	353	166	166		

- Molecule 15 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	a	90	Total	C	N	O	0	0
			372	192	90	90		

- Molecule 16 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	b	74	Total	C	N	O	0	0
			308	160	74	74		

- Molecule 17 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	c	79	Total 319	C 161	N 79	O 79	0	0

- Molecule 18 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	d	74	Total 305	C 157	N 74	O 74	0	0

- Molecule 19 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	e	83	Total 341	C 175	N 83	O 83	0	0

- Molecule 20 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
20	f	71	Total 293	C 151	N 71	O 71	0	0

- Molecule 21 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	g	82	Total 337	C 173	N 82	O 82	0	0

- Molecule 22 is [(2 {S},3 {S},4 {E},6 {S},7 {R},10 {R})-3,7-dimethyl-2-[(2 {E},4 {E},6 {R})-6-methyl-6-oxidanyl-7-[(2 {R},3 {R})-3-[(2 {R},3 {S})-3-oxidanylpentan-2-yl]oxiran-2-yl] hepta-2,4-dien-2-yl]-7,10-bis(oxidanyl)-12-oxidanylidene-1-oxacyclododec-4-en-6-yl] 4-cyclo heptylpiperazine-1-carboxylate (three-letter code: 9B0) (formula: $C_{40}H_{66}N_2O_9$) (labeled as "Ligand of Interest" by depositor).

9B0
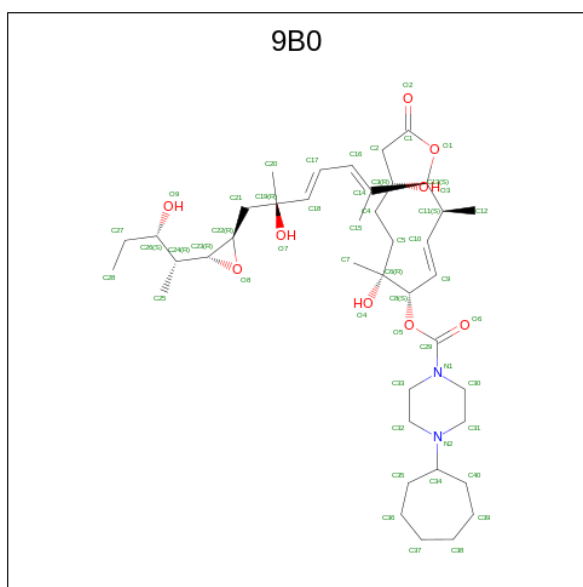
Mol	Chain	Residues	Atoms				AltConf
22	1	1	Total 51	C 40	N 2	O 9	0

- Molecule 23 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
23	6	3	Total 3	Zn 3	0
23	C	1	Total 1	Zn 1	0

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: U2 snRNA

- Molecule 2: Splicing factor 3B subunit 1

• Molecule 3: Splicing factor 3B subunit 2

9%

Chain 2:   17% • 80%

• Molecule 4: Splicing factor 3B subunit 3

Chain 3:

16%

73%

23%

• Molecule 5: Splicing factor 3B subunit 4

• Molecule 6: Splicing factor 3B subunit 5

• Molecule 7: PHD finger-like domain-containing protein 5A

Chain 6:

64% 14% 23%

MET ALA LYS HIS PRO D7 L8 R21 K25 C26 D27 G28 K29 Y36 V37 R38 C46 D47 E48 Y54 R57 C58 Q62 S67 D68 K73 N90 L91 GLY SER SER LYS THR ASP LEU PHE TYR GLU ARG LYS LYS TYR GLY PHE LYS LYS ARG

• Molecule 8: Splicing factor 3A subunit 1

16%

Chain A:

16% 84%

MET PRO ALA GLY PRO VAL GLN ALA VAL PRO PRO PRO PRO VAL PRO THR GLU GLU LYS GLN PRO THR GLU GLU ALA SER SER LYS GLU ASP SER ALA PRO SER LYS PRO VAL VAL GLY ILE ILE TYR PRO PRO GLU GLU VAL ARG ASN ILE VAL ASP LYS THR ALA SER PHE VAL

ALA ARG ASN GLY PRO GLU PHE GLU ALA ARG ILE ARG ASN GLU GLN ILE ASN ASN PRO LYS PHE ASN PHE LEU ASN PRO ASN ASP PRO TYR HIS ALA TYR TYR ARG HIS LYS VAL SER GLU PHE LYS GLY GLY LYS ALA GLN GLU GLU PRO SER ALA ILE PRO LYS VAL MET GLN GLN

GLN GLN THR THR GLN GLN GLN LEU PRO GLN LYS VAL GLN ALA GLN VAL ILE GLU GLN THR ILE VAL PRO LYS GLU GLU PRO PRO GLU PHE GLU PHE ILE ALA ASP PRO PRO SER ILE S160 A161 F162 D163 L164 D165 V166 V167 K168 L169 T170 A171 Q172 F173 V174 A175 R176 N177 G178 R179 Q180

F181 L182 T183 Q184 L185 M186 Q187 K188 E189 Q190 R191 N192 Y193 Q194 F195 D196 F197 L198 R199 P200 Q201 H202 S203 L204 F205 N206 Y207 F208 T209 K210 L211 V212 E213 Q214 Y215 T216 K217 I218 L219 I220 P221 P222 K223 G224 L225 F226 S227 K228 L229 K230 K231 E232 A233 E234 N235 P236 R237 E238 V239 L240

D241 Q242 V243 C244 Y245 R246 V247 E248 W249 A250 K251 F252 Q253 E254 R255 E256 R257 K258 K259 E260 E261 E262 E263 K264 E265 K266 E267 R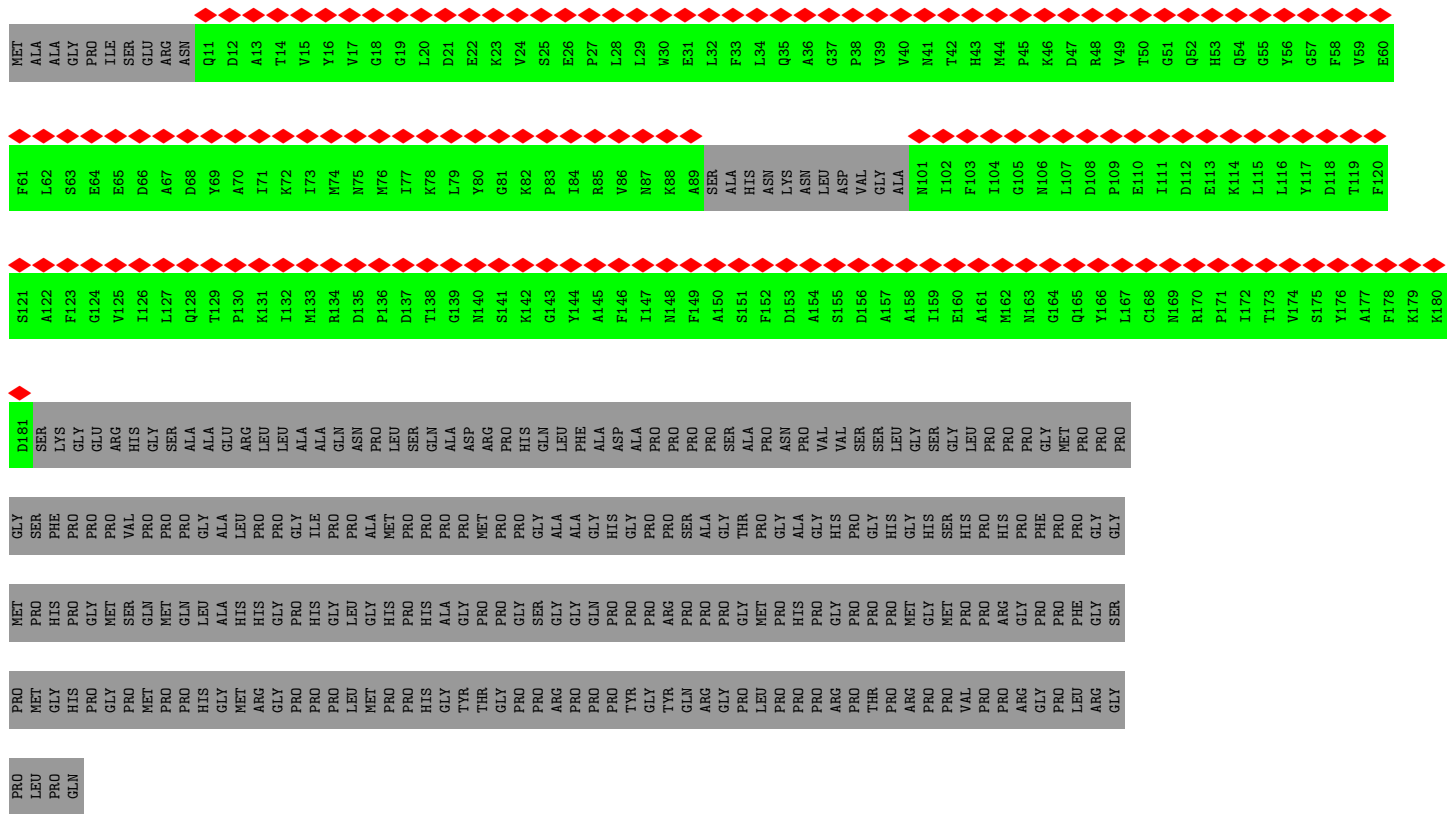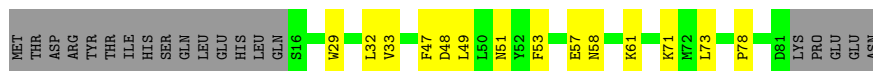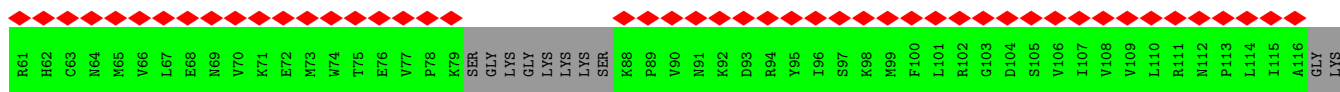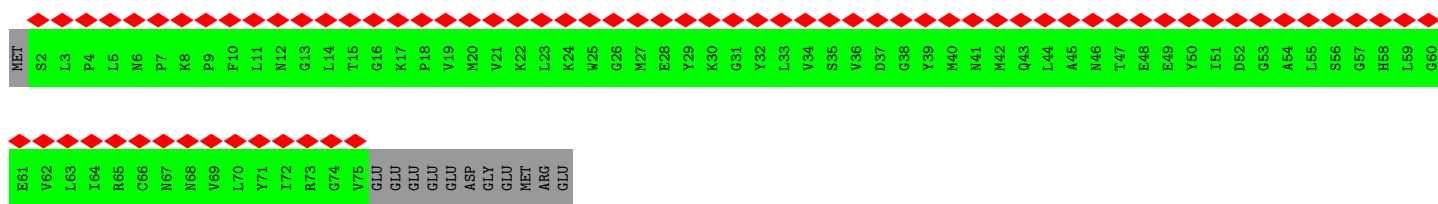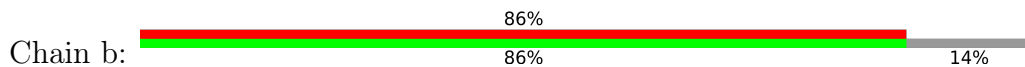268 V269 A270 Y271 A272 Q273 I274 D275 W276 H277 D278 F279 V280 V281 V282 GLU THR VAL ASP PHE GLN PRO ASN GLU GLN GLY ASN PHE PRO PRO THR THR

PRO GLU GLU GLY LEU GLY ALA ARG ILE LEU ILE GLN GLU GLU ARG TYR GLU LYS PHE GLY GLU GLU SER GLU GLU GLU VAL GLU MET GLU VAL GLU SER ASP GLU GLU ASP LYS GLN GLU GLN LYS ALA GLU GLU GLU PRO SER GLU LEU ASP ASP GLN THR GLN VAL VAL GLN LYS ASP MET ASP GLY GLY SER ASP

ASP GLU GLU GLY GLN ASP LYS VAL PRO PRO PRO GLU THR PRO MET PRO PRO PRO LEU PRO PRO THR PRO ASP GLN GLN VAL ILE VAL LYS ASP TYR ASP PRO LYS ALA SER LYS ALA PRO GLU TYR LEU VAL SER GLU ILE THR GLY GLU LYS ILE

PRO ALA SER LYS MET GLN HIS MET ARG ILE GLY LEU LEU ASP PRO ARG TRP LEU GLU GLU ARG ASP SER ILE ARG GLU LYS GLN SER ASP ASP VAL TYR ALA PRO GLY LEU ILE ASP GLU GLU SER LEU LYS LEU ALA GLU ARG ARG ASP ILE PHE GLY VAL GLU

GLU THR ALA ILE GLY LYS LYS ILE GLY GLU GLU ILE GLN LYS PRO GLU GLU LYS VAL THR TRP ASP GLY HIS SER GLY SER MET ALA ARG THR GLN ALA GLN ALA ASN ILE THR LEU GLU GLN GLN ILE GLU ALA ILE HIS LYS ALA LYS GLY LEU VAL PRO GLU GLU ASP ASP

THR LYS GLU LYS ILE GLY PRO SER LYS PRO ASN GLU ILE PRO GLN GLN PRO PRO PRO SER SER ALA THR ASN ILE PRO SER SER ALA ARG PRO ILE THR SER VAL PRO ARG PRO PRO THR MET ALA PRO ARG PRO PRO VAL MET ILE VAL ARG THR THR VAL VAL SER ALA VAL PRO VAL MET PRO ARG PRO

PRO MET ALA SER VAL VAL ARG LEU PRO PRO GLY SER VAL ILE ALA PRO MET PRO ILE HIS ALA PRO ARG ILE ASN VAL VAL PRO MET PRO SER ALA PRO ILE MET PRO SER ALA PRO ILE MET ALA PRO ARG PRO PRO MET ILE VAL PRO THR ALA PHE VAL PRO ALA PRO PRO VAL ALA PRO

VAL PRO ALA PRO ALA PRO MET PRO PRO VAL HIS PRO PRO PRO MET GLU ASP GLU THR PRO THR SER LYS LYS THR GLU SER SER VAL ILE LYS VAL HIS GLU ALA THR GLY MET PRO ALA GLY LYS GLN LYS LEU GLN TYR ASN MET GLN ASP LYS THR GLU TRP

LYS LEU ASN GLY VAL LEU LEU PHE THR LEU PRO LEU THR ASP GLN VAL SER VAL ILE LYS VAL ILE HIS GLU ALA THR GLY MET PRO ALA GLY LYS GLN LYS LEU GLN LEU GLN TYR ASN ALA ASN MET ALA ASN GLY ALA VAL

ILE HIS LEU LEU LEU LYS GLU ARG GLY GLY ARG LYS LYS

• Molecule 9: Splicing factor 3A subunit 2

Chain B:

22%

22%

78%

• Molecule 10: Splicing factor 3A subunit 3

Chain C:

65%

77%

6%

16%

E361 E362 GLN ILE SER GLU SER GLU SER GLU ASP GLU GLU ASN GLU ILE ILE TYR ASN PRO LYS ASN LEU PRO LEU GLY TRP ASP GLY LYS PRO ILE P393 Y394 W395 L396 Y397 K398 L399 H400 G401 L402 M403 I404 N405 Y406 C411 C412 N413 Y414 T415 Y416 R417 H425 F426 A427 E428 W429

R430 C437 L438 T452 S457 K461 K466 A467 S468 E469 R470 W471 Q472 T475 E476 E477 D481 S482 S483 G484 N485 V486 K489 K490 D494 Q498 G499 L500 L501

• Molecule 11: HIV Tat-specific factor 1

Chain D:

13%

22% • 74%

MET SER GLY THR ASN LEU ASP GLY ASN GLU PHE GLU GLN GLN LEU ARG MET GLU GLU LEU TYR ASP GLY LYS ASP GLY THR GLN THR ASP ALA GLY PRO SER LEU GLY GLN PRO ASP THR TYR TRP ASP LYS ALA TRP PHE

PRO LYS ILE GLU ASP PHE ILE ALA THR TYR GLN ALA ASN TYR PHE GLY PHE SER ASN ASP GLY ALA SER SER THR ALA ASN VAL GLU ASP VAL HIS ALA ARG THR ALA GLU GLU PRO GLN GLU LYS ALA PRO GLU PRO THR ASP ARG LYS LYS GLY GLU LYS ARG LYS ALA

GLU SER GLY TRP PHE HIS V127 E128 E129 D130 Y136 V137 S138 D143 D147 I150 F156 M160 R161 D162 T165 E166 E167 F168 L172 Q177 G178 N179 C187 L201 I206 R207 G208 Y209 H212 V213 E214 V215 A216 K217 L220 LYS GLY GLU TYR ASP

ALA SER LYS LYS LYS LYS CYS LYS D235 Y236 K237 K238 K239 S241 L240 M242 K245 Q246 L247 D248 W249 R250 R253 R254 A255 G256 PRO SER MET ARG HIS GLU R264 V265 V266 I267 I268 K269 N270 M271 F272 H273 P274 M275 D276 F277 E278 D279 D280 P281 L282 V283 L284 N285 E286 I287

R288 E289 D290 L291 R292 V293 E294 C295 S296 K297 F298 G299 Q300 I301 R302 K303 L304 L305 L306 F307 D308 R309 H310 P311 D312 G313 V314 A315 S316 V317 S318 F319 R320 D321 P322 E323 E324 A325 D326 Y327 C328 I329 Q330 T331 L332 D333 G334 R335 W336 F337 G338 G339 R340 Q341 I342 T343 A344 Q345 A346 W347

ASP GLY THR THR ASP TYR GLN VAL GLU GLU GLN GLY THR SER ARG GLU GLU GLU ARG GLU ARG GLY LEU ARG ARG GLY TRP GLU ALA PHE LEU ASN ALA PRO GLU GLU GLN ALA ASN ARG GLY LEU ARG SER ASP SER VAL THR ALA SER GLU GLU ALA ALA GLY ALA GLY PRO SER ARG ALA ARG ALA HIS PHE SER HIS PRO SER

THR SER LYS MET ASN ALA GLU GLU GLU GLN THR ALA THR GLY MET ALA PHE GLU GLU PRO ILE ASP LYS LYS PHE GLU LYS THR GLU ASP GLY GLU GLU GLY GLY GLU PHE GLU GLU GLY GLY ALA SER ASP CYS SER GLU LYS GLN LEU THR LEU LYS ASN ASP ASP CYS GLU GLU SER PRO GLU GLY LYS GLU SER

GLU GLU GLY CYS PRO LYS ARG GLY PHE GLU GLY GLY SER SER CYS SER GLN GLU SER GLU GLU GLY GLY ASN PRO VAL ARG GLY SER GLU GLU ASP PRO LYS GLU SER LYS LYS THR LEU LYS ASN ASP CYS GLU GLU ASN GLY GLU ALA LYS GLU GLY GLU SER ASP ASP LEU ASN LYS

GLU SER GLU GLU GLU VAL PRO THR LYS GLU GLU SER CYS SER GLN GLU ASP SER SER GLU ASP LYS GLU ASP GLU PHE GLU ASP GLY ASP LYS SER ARG GLU PHE GLU ASP SER ALA ASP LYS LEU GLU GLU GLU GLU SER GLU GLY ASP ASP LEU HIS GLU ASN VAL

LEU ASP LYS GLU LEU GLU GLU ASN ASP SER GLU ASN ASP GLU PHE GLU ASP SER GLU GLY LYS VAL LEU ASP GLU GLY GLY SER GLU GLU ASP ALA SER GLU ASP LYS ASP THR TYR GLU LYS VAL PHE ASP GLU ASP SER ASP LYS GLU LEU ASP

GLU GLU TYR ALA ASP GLU LYS GLY LEU GLU ALA ALA ASP LYS LYS ALA GLU GLU LYS ASP ALA GLU ASP LEU PHE GLU GLU ASP SER SER ASN GLU LYS LEU PHE

ASP GLU GLU GLU ASP SER SER GLU LYS LEU PHE ASP ASP SER ASP ARG GLY THR LEU GLY PHE GLY SER VAL GLU GLU GLY PRO LEU SER THR GLY SER SER PHE ILE LEU SER ASP ASP ASP ASP ILE

• Molecule 12: RNA helicase

Chain E:

41%

41% • 57%

MET ALA SER ASP TYR LYS ASP ASP ASP LYS ALA SER ASP GLU VAL ASP ALA GLY THR MET GLY ARG GLU ARG HIS TYR ARG LYS ARG SER ALA SER ARG SER ARG GLY SER ALA SER ARG SER PRO SER ASP LYS ARG SER LYS ARG SER ARG GLY ASP ASP ARG ARG SER

ARG SER ARG ASP ASP ARG ARG ARG GLU SER ARG SER LYS LYS THR GLU ASP LYS ARG SER ARG SER LYS GLU LYS THR ASP GLY GLY GLU SER LYS GLU LYS LYS LYS ASP LYS ARG ARG SER ARG SER ARG ARG GLU ARG ASP ARG SER ARG SER ARG ARG ARG SER ARG SER ARG ASP ARG ARG ARG SER ARG SER ARG ARG GLY ARG

ARG SER ARG SER SER PRO GLY ASN ASN LYS SER LYS LYS THR GLU ASN ARG SER ARG SER LYS GLU LYS THR ASP GLY GLY GLU SER LYS GLU LYS LYS LYS ASP LYS ASP LYS GLU ASP LYS ALA GLY ASN PHE D152 Q153 N154 K155 L156 E157 M160 R161

E165 R166 V167 E168 K169 E172 E173 Q174 R175 K176 K177 A178 MET GLU ASN ILE GLU GLY GLY GLY LEU LYS LYS GLU ILE GLU GLU MET LYS GLN G195 K196 K197 D203 D204 D205 D206 E207 D208 D209 PRO ALA ALA GLU LYS LYS GLY GLU ASN GLU MET GLU LYS GLY GLU GLU L225 D226 P227 L228

D229 E233 K236 E237 E238 V239 K240 K241 F242 N243 MET ARG SER VAL LYS GLY GLY GLY ASN GLU LYS LYS GLU LEU LEU GLU GLU PRO VAL ASP HIS GLY LYS LYS ILE GLU GLU TYR GLU PRO PHE ARG LYS ASN PHE TYR VAL GLU VAL PRO GLU LEU ALA LYS MET SER GLN GLU GLN ASP ALA

MET GLU TYR SER GLU SER GLU GLU GLU VAL ASP LEU GLN THR ALA LEU THR GLY TYR GLN THR LYS GLN ARG GLU LEU GLU GLU PRO VAL ASP HIS GLY LYS LYS ILE GLU GLU TYR GLU PRO PHE ARG LYS ASN PHE TYR VAL GLU VAL PRO GLU LEU ALA LYS MET SER GLN GLU GLU V350 N351

V352 F353 R354 L355 E356 M357 E358 G359 I360 T361 V362 K363 G364 K365 G366 C367 P368 K369 P370 I371 K372 S373 W374 V375 Q376 C377 G378 I379 S380 M381 K382 I383 L384 N385 S386 L387 K388 K389 H390 G391 Y392 E393 K394 P395 T396 P397 I398 Q399 T400 Q401 A402 I403 P404 A405 I406 M407 S408 G409 R410 D411

L412 I413 G414 I415 A416 K417 T418 G419 S420 G421 K422 T423 I424 A425 F426 L427 L428 P429 M430 F431 R432 H433 I434 M435 D436 Q437 R438 S439 L440 E441 E442 G443 E444 G445 P446 I447 A448 V449 I450 M451 T452 P453 T454 R455 E456 L457 A458 L459 Q460 I461 T462 K463 E464 C465 K466 K467 F468 S469 K470 T471

L472 G473 L474 R475 V476 V477 C478 V479 Y480 G481 G482 T483 G484 I485 S486 E487 Q488 I489 A490 E491 L492 K493 R494 G495 A496 E497 I498 I499 V500 C501 T502 P503 G504 R505 M506 I507 D508 M509 L510 A511 A512 N513 S514 G515 R516 V517 T518 N519 L520 R521 R522 V523 T524 Y525 V526 V527 L528 D529 E530 A531

D532 R533 M534 F535 D536 M537 G538 F539 E540 P541 Q542 V543 M544 R545 I546 V547 D548 N549 V550 R551 P552 D553 R554 Q555 T556 V557 M558 F559 S560 A561 T562 F563 P564 R565 A566 M567 E568 A569 L570 A571 R572 R573 I574 L575 S576 K577 P578 I579 E580 V581 Q582 V583 G584 G585 R586 S587 V588 V589 C590 S591

D592 V593 E594 Q595 Q596 V597 I598 V599 I600 E601 E602 E603 K604 K605 F606 L607 K608 L609 L610 E611 L612 L613 G614 H615 Y616 Q617 E618 S619 G620 S621 V622 I623 I624 F625 V626 D627 K628 Q629 E630 H631 A632 D833 G634 L635 L636 K637 D638 L639 M640 R641 A642 S643 Y644 P645 C646 M647 S648 L649 H650 G651

G652 I653 D654 Q655 Y656 D657 R658 D659 S660 I661 I662 N663 D664 F665 K666 N667 G668 T669 C670 K671 L672 L673 V674 A675 T676 S677 V678 A679 A680 R681 G682 L683 D684 V685 K686 H687 L688 I689 L690 V691 V692 N693 Y694 S695 C696 P697 N698 H699 Y700 E701 D702 Y703 V704 H705 R706 A707 G708 R709 T710 G711

R712 A713 G714 N715 K716 G717 Y718 A719 Y720 T721 F722 I723 T724 E725 D726 Q727 A728 R729 Y730 A731 G732 D733 I734 I735 K736 A737 L738 E739 L740 S741 G742 T743 ALA VAL PRO PRO ASP GLU LEU GLU LYS LEU TRP SER ASP PHE LYS ASP GLN GLN LYS ALA GLU GLY LYS ILE ILE LYS LYS SER SER

GLY PHE SER GLY LYS GLY PHE LYS PHE ASP GLU THR GLU GLN ALA LEU LEU ASN ALA LYS LEU GLU ILE ARG LYS LYS LEU GLN LYS ALA ALA LEU GLY LEU GLN ASP SER ASP ASP ASP ALA ALA VAL ASP ILE ASP GLU GLU GLN ILE GLU GLU SER MET PHE ASN SER LYS LYS ARG VAL LYS ASP MET ALA

ALA PRO GLY THR SER SER VAL PRO ALA PRO ALA THR ALA GLY LYS ASN ALA LYS LEU GLU ILE ALA LYS LEU ARG LEU ARG ILE ASN ALA GLN LYS ASN LEU GLY LYS LEU GLY GLN ASN ALA ILE LEU ARG GLY THR THR ILE LEU ALA PRO THR

VAL SER ALA LYS THR ILE ALA GLU GLN GLU LEU ALA GLU GLU LYS ILE ASN ALA LYS LEU ASN TYR VAL GLU GLU GLU GLU LEU GLU ILE ASN ASP PHE PRO GLN THR ALA ARG TRP LYS VAL THR

SER LYS GLU ALA LEU GLN ARG ILE SER GLU GLU TYR SER GLU ALA ALA ILE ILE THR ILE ARG GLY THR TYR PHE PRO PRO GLY LYS GLU PRO LYS GLU GLY GLY ARG LYS ILE TYR LEU ALA ILE GLU GLU SER ALA ASN GLU LEU ALA VAL GLN LYS ALA LYS ALA GLU ILE THR ARG LEU ILE LYS

GLU GLU LEU ILE ARG LEU GLN ASN SER TYR GLN PRO THR ASN LYS GLY ARG TYR LYS VAL LEU

• Molecule 13: U2 small nuclear ribonucleoprotein A’

Chain F:
64%
64%   36%

MET V2 K3 L4 T5 A6 E7 L8 I9 E10 Q11 A12 A13 Q14 Y15 T16 N17 A18 V19 R20 D21 R22 E23 L24 D25 L26 R27 G28 Y29 K30 I31 P32 V33 I34 E35 N36 L37 G38 A39 T40 L41 D42 Q43 F44 D45 A46 I47 D48 F49 S50 D51 N52 E53 I54 R55 K56 L57 D58 G59 F60

P61 L62 L63 R64 R65 L66 K67 T68 L69 L70 V71 N72 N73 N74 R75 I76 C77 R78 I79 G80 E81 G82 L83 D84 Q85 A86 L87 P88 C89 L90 T91 E92 L93 I94 L95 T96 N97 N98 S99 L100 V101 E102 L103 G104 D105 L106 D107 P108 L109 A110 S111 L112 K113 S114 L115 T116 Y117 L118 S119 I120

L121 R122 N123 P124 V125 T126 N127 K128 K129 H130 Y131 R132 L133 Y134 V135 I136 Y137 K138 V139 P140 Q141 V142 R143 V144 L145 D146 F147 Q148 K149 V150 K151 L152 K153 E154 R155 Q156 E157 A158 E159 K160 M161 F162 K163 GLY LYS ARG GLY ALA GLN LEU ALA LYS ASP ILE ALA ARG ARG SER LYS THR

PHE ASN PRO GLY ALA GLY LEU PRO THR ASP LYS LYS GLY GLY PRO SER PRO GLY ASP VAL GLU ALA ILE LYS ASN ALA ILE ALA ASN ALA SER THR LEU ALA GLU VAL GLU ARG LEU LYS GLY LEU LEU GLN SER GLY GLN ILE PRO GLY ARG GLU GLU ARG SER GLY PRO THR ASP

ASP GLY GLU GLU GLU MET GLU GLU ASP THR VAL THR ASN GLY SER

• Molecule 14: U2 small nuclear ribonucleoprotein B”

Chain G:
74%
73%   26%

MET ASP I6 R7 P8 N9 H10 T11 I12 Y13 I14 N15 N16 M17 N18 D19 K20 I21 K22 K23 E24 E25 L26 K27 R28 S29 L30 Y31 A32 L33 F34 S35 Q36 F37 G38 H39 V40 V41 D42 I43 V44 A45 L46 K47 T48 M49 K50 M51 R52 G53 Q54 A55 F56 V57 I58 F59 K60 E61 L62 G63

S64 S65 T66 N67 A68 L69 R70 Q71 L72 Q73 G74 F75 P76 F77 Y78 G79 K80 P81 M82 R83 I84 Q85 Y86 A87 K88 T89 D90 S91 D92 I93 I94 S95 K96 M97 R98 G99 THR PHE ALA ASP LYS GLU LYS LYS LYS GLU LYS LYS LYS ALA LYS THR VAL GLU GLN THR ALA THR THR

ASN LYS LYS PRO GLY GLN GLY THR PRO ASN SER ALA ASN THR GLN GLY ASN SER THR PRO ASN PRO GLN VAL PRO ASP TYR PRO P149 N150 Y151 I152 L153 F154 L155 N156 N157 L158 P159 E160 E161 T162 N163 E164 M165 M166 L167 S168 M169 L170 F171 N172 Q173 F174 P175 G176 F177 K178 E179 V180

R181 L182 VAL PRO GLY ARG HIS D188 I189 A190 F191 V192 E193 F194 E195 N196 D197 G198 Q199 A200 G201 A202 A203 R204 D205 A206 L207 Q208 G209 F210 K211 I212 T213 P214 S215 H216 A217 M218 K219 I220 T221 Y222 A223 K224 K225

• Molecule 15: Small nuclear ribonucleoprotein Sm D2

Chain a:
76%
76%   24%

MET SER LEU LEU ASN LYS PRO LYS SER GLU MET THR PRO GLU GLU LEU GLN LYS R19 E20 E21 E22 E23 F24 N25 T26 G27 P28 L29 S30 V31 L32 T33 Q34 S35 V36 K37 N38 N39 T40 Q41 V42 L43 I44 N45 C46 R47 N48 N49 K50 K51 L52 L53 G54 R55 V56 K57 A58 F59 D60

R61 H62 C63 N64 M65 V66 L67 E68 N69 V70 K71 E72 M73 W74 T75 E76 V77 P78 K79 SER GLY LYS GLY LYS LYS SER K88 P89 V90 N91 K92 D93 R94 Y95 I96 S97 K98 M99 F100 L101 R102 G103 D104 S105 V106 I107 V108 V109 L110 R111 N112 P113 L114 I115 A116 GLY LYS

• Molecule 16: Small nuclear ribonucleoprotein F

86%

Chain b:

86%   14%

MET S2 L3 P4 L5 N6 P7 K8 P9 F10 L11 N12 G13 L14 T15 G16 K17 P18 V19 M20 V21 K22 L23 K24 W25 G26 M27 E28 Y29 K30 G31 Y32 L33 V34 S35 V36 D37 G38 Y39 M40 N41 M42 Q43 L44 A45 N46 T47 E48 E49 Y50 I51 D52 G53 A54 L55 S56 G57 H58 L59 G60

E61 V62 L63 I64 R65 C66 N67 N68 V69 L70 Y71 I72 R73 G74 V75 GLU GLU GLU GLU GLU ASP GLY GLU MET ARG GLU

• Molecule 17: Small nuclear ribonucleoprotein E

86%

Chain c:

86%   14%

MET ALA TYR ARG GLY GLN GLY GLN LYS VAL GLN LYS VAL M14 V15 Q16 P17 I18 N19 L20 I21 F22 R23 Y24 L25 Q26 N27 R28 S29 R30 I31 Q32 V33 W34 L35 Y36 E37 Q38 V39 N40 M41 R42 I43 E44 G45 C46 I47 I48 G49 F50 D51 E52 Y53 M54 N55 L56 V57 L58 D59 D60

A61 E62 E63 I64 H65 S66 K67 T68 K69 S70 R71 K72 Q73 L74 G75 R76 I77 M78 L79 K80 G81 D82 N83 I84 T85 L86 L87 Q88 S89 V90 S91 N92

• Molecule 18: Small nuclear ribonucleoprotein G

97%

Chain d:

97%   .

MET SER K3 A4 H5 P6 P7 E8 L9 K10 K11 F12 M13 D14 K15 K16 L17 S18 L19 K20 L21 N22 G23 G24 R25 H26 V27 Q28 G29 I30 L31 R32 G33 F34 D35 P36 F37 M38 N39 L40 V41 I42 D43 E44 C45 V46 E47 M48 A49 T50 S51 G52 Q53 Q54 N55 N56 I57 G58 M59 V60

V61 I62 R63 G64 N65 S66 I67 I68 M69 L70 E71 A72 L73 E74 R75 V76

• Molecule 19: Small nuclear ribonucleoprotein Sm D3

66%

Chain e:

66%   34%

MET S2 I3 G4 V5 P6 I7 K8 V9 L10 H11 E12 A13 E14 G15 H16 I17 V18 T19 C20 E21 T22 N23 T24 G25 E26 V27 Y28 R29 G30 K31 L32 I33 E34 A35 E36 D37 N38 M39 N40 C41 Q42 M43 S44 N45 I46 T47 V48 T49 Y50 R51 D52 G53 R54 V55 A56 Q57 L58 E59 Q60

V61 Y62 I63 R64 G65 S66 K67 I68 R69 F70 L71 I72 L73 P74 D75 M76 L77 K78 N79 A80 P81 M82 L83 K84 SER MET LYS ASN LYS ASN GLN GLY SER GLY ALA GLY ARG GLY LYS ALA ALA ILE LEU LYS ALA GLN VAL ALA ALA ARG GLY ARG GLY ARG GLY MET GLY ARG GLY ASN

ILE PHE GLN LYS ARG

• Molecule 20: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain f:

30%

30%   70%

MET THR VAL G4 K5 S6 S7 K8 M9 L10 Q11 H12 I13 D14 Y15 R16 M17 R18 C19 I20 L21 Q22 D23 G24 R25 I26 F27 I28 G29 T30 F31 K32 A33 F34 D35 K36 H37 M38 N39 L40 I41 L42 C43 D44 C45 D46 E47 F48 R49 LYS ILE LYS PRO LYS ASN SER LYS GLN ALA GLU

ARG GLU E63 K64 R65 V66 L67 G68 L69 V70 L71 L72 R73 G74 E75 N76 L77 V78 S79 M80 T81 V82 E83 G84 P85 P86 P87 LYS ASP THR GLY ILE ALA ARG VAL PRO LEU ALA GLY ALA ALA GLY GLY PRO GLY ILE GLY ARG ALA ALA GLY ARG ILE PRO ALA GLY VAL PRO MET

PRO GLN ALA PRO ALA GLY LEU ALA GLY PRO VAL ARG GLY VAL GLY PRO SER GLN GLN VAL MET THR PRO GLN GLY ARG GLY THR VAL ALA ALA ALA ALA ALA THR ALA SER ILE ALA GLY ALA PRO THR GLN TYR PRO PRO GLY ARG GLY PRO PRO PRO MET GLY

ARG GLY ALA PRO PRO PRO GLY MET GLY GLY PRO PRO GLY MET ARG PRO PRO MET GLY PRO PRO MET GLY ILE PRO PRO GLY ARG GLY THR PRO PRO MET GLY MET PRO PRO PRO GLY MET ARG PRO PRO PRO PRO GLY MET ARG PRO PRO ARG PRO

• Molecule 21: Small nuclear ribonucleoprotein Sm D1

Chain g:

69%

69%   31%

M1 K2 L3 V4 R5 F6 L7 M8 K9 L10 S11 H12 E13 T14 V15 T16 I17 E18 L19 K20 N21 G22 T23 Q24 V25 H26 G27 T28 I29 T30 G31 V32 D33 V34 S35 M36 N37 T38 H39 L40 K41 A42 V43 K44 M45 T46 L47 K48 N49 R50 E51 P52 V53 Q54 L55 E56 T57 L58 S59 I60

R61 G62 N63 N64 I65 R66 Y67 F68 I69 L70 P71 D72 S73 L74 P75 L76 D77 T78 L79 L80 V81 D82 VAL GLU PRO LYS VAL LYS SER LYS LYS ARG GLU GLY ALA VAL ALA GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY PRO ARG ARG

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	485418	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.257	Depositor
Minimum map value	-0.135	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	347.84, 347.84, 347.84	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.087, 1.087, 1.087	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 9B0

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	H	0.41	2/3507 (0.1%)	0.95	17/5454 (0.3%)
2	1	0.27	0/6609	0.45	1/8946 (0.0%)
3	2	0.25	0/1248	0.45	0/1672
4	3	0.27	0/9435	0.49	1/12802 (0.0%)
5	4	0.24	0/670	0.45	0/850
6	5	0.48	1/555 (0.2%)	0.52	0/750
7	6	0.25	0/653	0.45	0/877
8	A	0.24	0/507	0.41	0/640
9	B	0.24	0/426	0.50	0/543
10	C	0.25	0/2133	0.45	0/2773
11	D	0.28	0/1325	0.47	0/1749
12	E	0.41	1/2171 (0.0%)	0.52	0/2779
13	F	0.24	0/671	0.49	0/849
14	G	0.23	0/689	0.46	0/869
15	a	0.24	0/374	0.50	0/472
16	b	0.25	0/311	0.51	0/395
17	c	0.24	0/319	0.49	0/399
18	d	0.25	0/307	0.51	0/388
19	e	0.25	0/343	0.51	0/433
20	f	0.24	0/294	0.49	0/370
21	g	0.23	0/339	0.51	0/428
All	All	0.30	4/32886 (0.0%)	0.56	19/44438 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	H	0	1
10	C	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	5	78	PRO	C-N	8.67	1.50	1.34
1	H	45	C	O3'-P	-7.13	1.52	1.61
1	H	55	U	P-OP2	5.57	1.58	1.49
12	E	178	ALA	C-O	5.12	1.33	1.23

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	28	C	N1-C1'-C2'	-11.54	98.99	114.00
1	H	120	A	C6-N1-C2	-8.05	113.77	118.60
1	H	42	G	C4'-C3'-O3'	7.34	127.69	113.00
1	H	27	U	C3'-C2'-C1'	7.02	107.11	101.50
1	H	55	U	OP1-P-OP2	6.85	129.88	119.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
10	C	393	PRO	Mainchain
1	H	43	U	Sidechain

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	H	3142	0	1594	60	0
2	1	6487	0	6700	136	0
3	2	1210	0	1010	27	0
4	3	9247	0	9179	161	0
5	4	664	0	230	0	0
6	5	539	0	505	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	6	645	0	626	9	0
8	A	504	0	150	0	0
9	B	421	0	148	0	0
10	C	2103	0	1108	47	0
11	D	1308	0	1061	25	0
12	E	2153	0	1054	27	0
13	F	666	0	209	0	0
14	G	685	0	226	7	0
15	a	372	0	118	0	0
16	b	308	0	115	0	0
17	c	319	0	92	0	0
18	d	305	0	105	0	0
19	e	341	0	112	0	0
20	f	293	0	99	0	0
21	g	337	0	110	0	0
22	1	51	0	0	9	0
23	6	3	0	0	0	0
23	C	1	0	0	0	0
All	All	32104	0	24551	454	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 8.

The worst 5 of 454 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:25:G:H1'	10:C:399:LEU:CD2	1.33	1.55
12:E:226:ASP:OD1	12:E:227:PRO:HD2	1.18	1.31
10:C:396:LEU:HD12	10:C:400:HIS:HD2	0.98	1.13
1:H:25:G:C1'	10:C:399:LEU:CD2	2.25	1.12
10:C:396:LEU:HD12	10:C:400:HIS:CD2	1.88	1.08

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	
2	1	813/1304 (62%)	782 (96%)	31 (4%)	0	100	100
3	2	172/895 (19%)	156 (91%)	15 (9%)	1 (1%)	25	43
4	3	1168/1217 (96%)	1121 (96%)	46 (4%)	1 (0%)	51	73
5	4	156/424 (37%)	153 (98%)	3 (2%)	0	100	100
6	5	64/86 (74%)	63 (98%)	1 (2%)	0	100	100
7	6	83/110 (76%)	79 (95%)	4 (5%)	0	100	100
8	A	121/793 (15%)	121 (100%)	0	0	100	100
9	B	96/464 (21%)	94 (98%)	2 (2%)	0	100	100
10	C	416/501 (83%)	398 (96%)	17 (4%)	1 (0%)	47	68
11	D	198/755 (26%)	182 (92%)	16 (8%)	0	100	100
12	E	447/1052 (42%)	431 (96%)	15 (3%)	1 (0%)	47	68
13	F	160/255 (63%)	160 (100%)	0	0	100	100
14	G	160/225 (71%)	158 (99%)	2 (1%)	0	100	100
15	a	86/118 (73%)	84 (98%)	2 (2%)	0	100	100
16	b	72/86 (84%)	71 (99%)	1 (1%)	0	100	100
17	c	77/92 (84%)	77 (100%)	0	0	100	100
18	d	72/76 (95%)	72 (100%)	0	0	100	100
19	e	81/126 (64%)	79 (98%)	2 (2%)	0	100	100
20	f	67/240 (28%)	67 (100%)	0	0	100	100
21	g	80/119 (67%)	79 (99%)	1 (1%)	0	100	100
All	All	4589/8938 (51%)	4427 (96%)	158 (3%)	4 (0%)	54	73

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
10	C	412	GLY
3	2	496	ASN
12	E	585	GLY
4	3	997	GLY

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	
2	1	701/1104 (64%)	666 (95%)	35 (5%)	24	46
3	2	104/776 (13%)	98 (94%)	6 (6%)	20	38
4	3	1022/1051 (97%)	963 (94%)	59 (6%)	20	38
5	4	8/336 (2%)	8 (100%)	0	100	100
6	5	57/77 (74%)	56 (98%)	1 (2%)	59	81
7	6	73/95 (77%)	72 (99%)	1 (1%)	67	86
8	A	4/709 (1%)	4 (100%)	0	100	100
9	B	7/382 (2%)	7 (100%)	0	100	100
10	C	80/446 (18%)	69 (86%)	11 (14%)	3	6
11	D	108/661 (16%)	100 (93%)	8 (7%)	13	27
12	E	73/909 (8%)	66 (90%)	7 (10%)	8	16
13	F	6/218 (3%)	6 (100%)	0	100	100
14	G	7/195 (4%)	7 (100%)	0	100	100
15	a	4/110 (4%)	4 (100%)	0	100	100
16	b	4/74 (5%)	4 (100%)	0	100	100
17	c	1/84 (1%)	1 (100%)	0	100	100
18	d	3/66 (4%)	3 (100%)	0	100	100
19	e	3/101 (3%)	3 (100%)	0	100	100
20	f	3/177 (2%)	3 (100%)	0	100	100
21	g	3/101 (3%)	3 (100%)	0	100	100
All	All	2271/7672 (30%)	2143 (94%)	128 (6%)	25	40

5 of 128 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	D	187	CYS
11	D	237	LYS
4	3	316	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	3	298	MET
12	E	157	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
4	3	805	ASN
10	C	400	HIS
4	3	870	ASN
4	3	983	ASN
11	D	179	ASN

5.3.3   RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	H	143/188 (76%)	27 (18%)	0

5 of 27 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	H	22	U
1	H	23	A
1	H	24	A
1	H	43	U
1	H	47	U

There are no RNA pucker outliers to report.

5.4   Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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	#$\|Z\| > 2$
22	9B0	1	1500	-	52,54,54	0.73	1 (1%)	61,77,77	0.80	2 (3%)

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
22	9B0	1	1500	-	-	25/67/91/91	0/3/4/4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
22	1	1500	9B0	C17-C16	-3.63	1.32	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
22	1	1500	9B0	C17-C16-C14	2.91	131.59	127.32
22	1	1500	9B0	O9-C26-C27	-2.47	104.86	109.35

There are no chirality outliers.

5 of 25 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
22	1	1500	9B0	O5-C29-N1-C33
22	1	1500	9B0	O5-C29-N1-C30
22	1	1500	9B0	N1-C29-O5-C8
22	1	1500	9B0	O5-C8-C9-C10
22	1	1500	9B0	C6-C8-C9-C10

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
22	1	1500	9B0	9	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand 9B0 1 1500

Bond lengths

Bond angles

Torsions

Rings
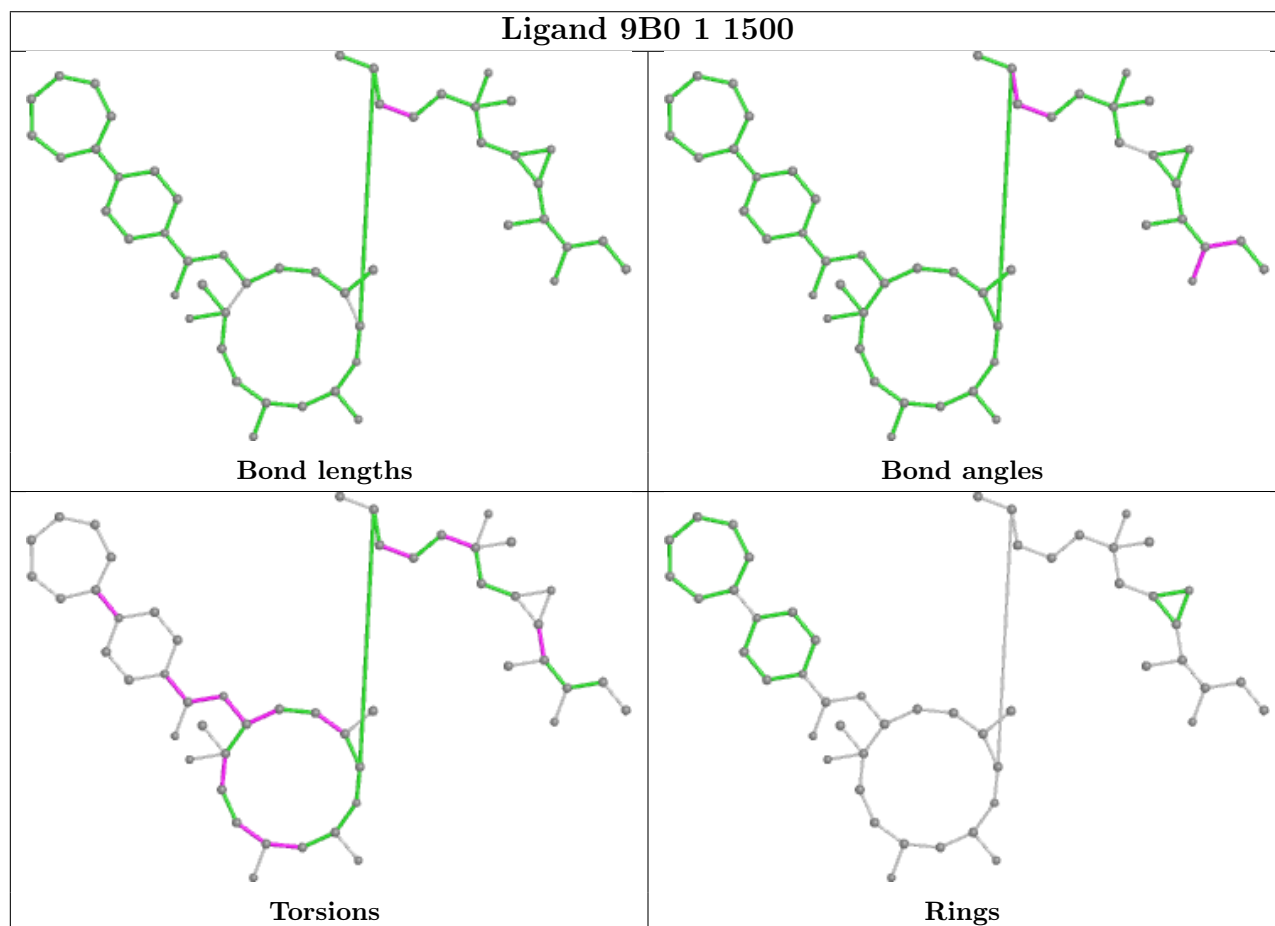
5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6   Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-31334. These allow visual inspection of the internal detail of the map and identification of artifacts.

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

6.1   Orthogonal projections ⓘ

6.1.1   Primary map

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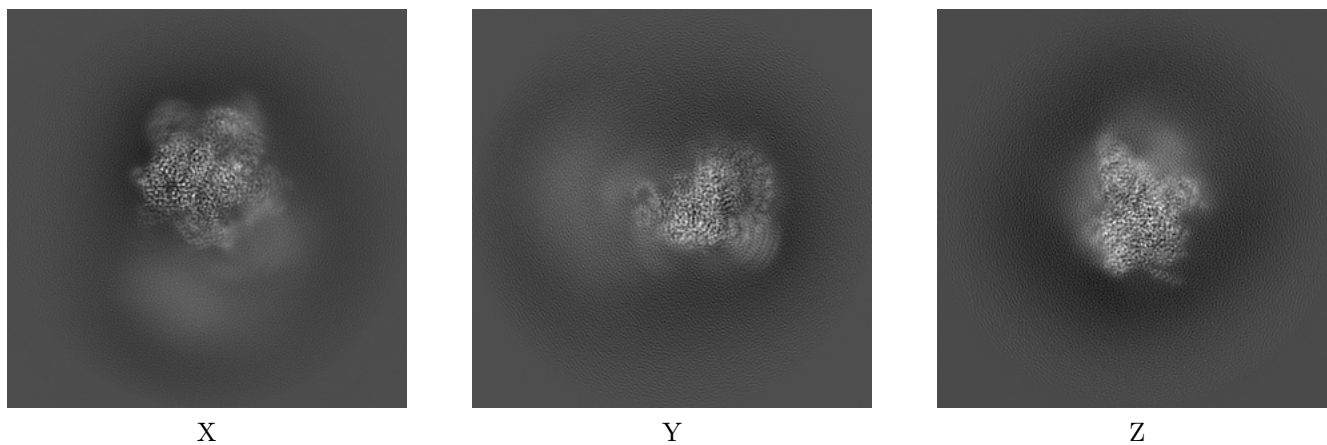
The images above show the map projected in three orthogonal directions.

6.2   Central slices ⓘ

6.2.1   Primary map

X Index: 160   Y Index: 160   Z Index: 160
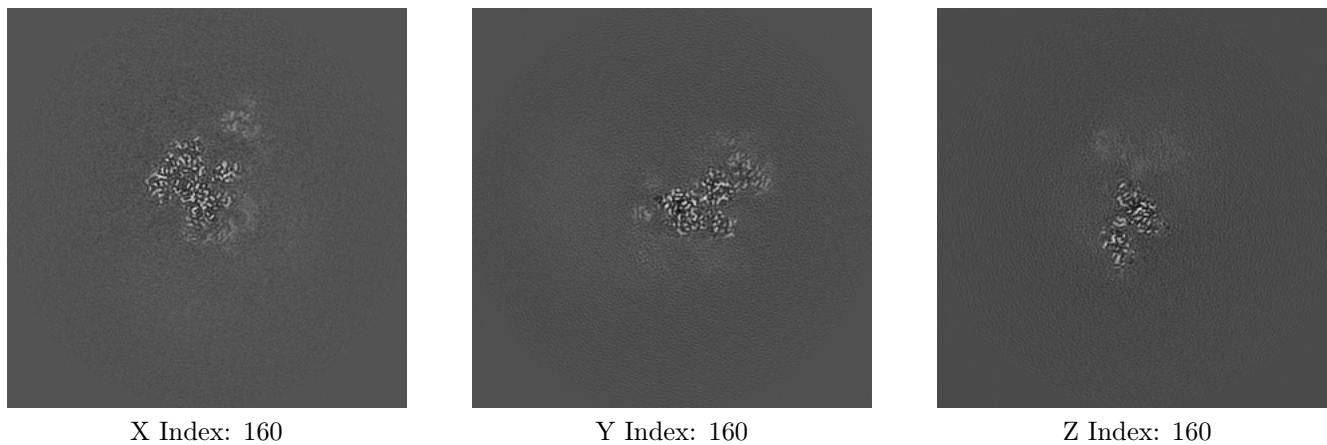
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ
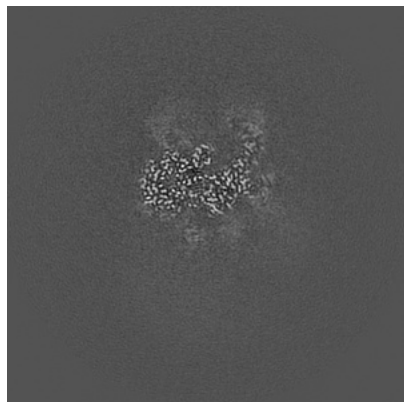
6.3.1 Primary map
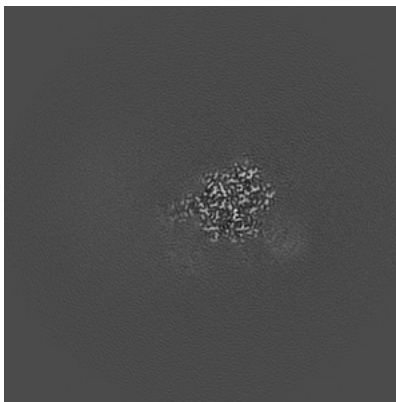
| X Index: 147 | Y Index: 148 | Z Index: 191 |
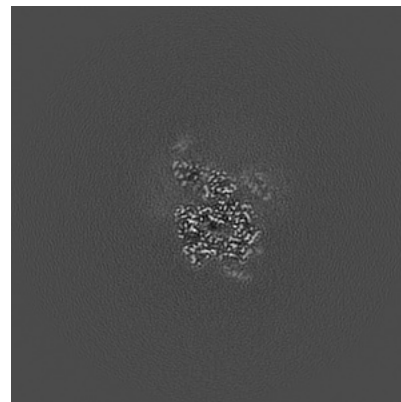
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map

| X | Y | Z |
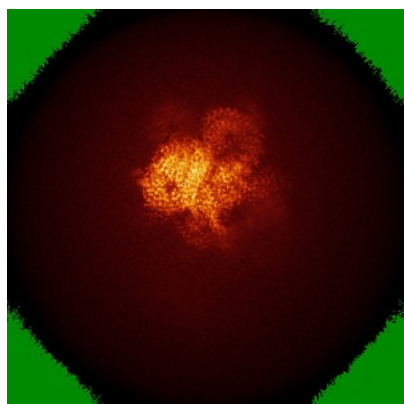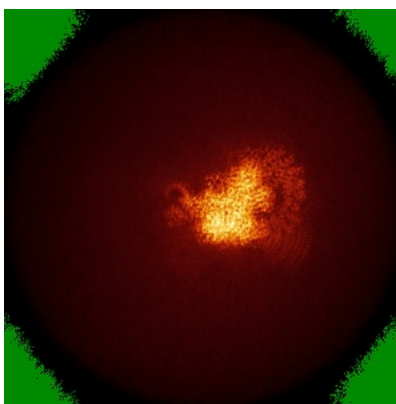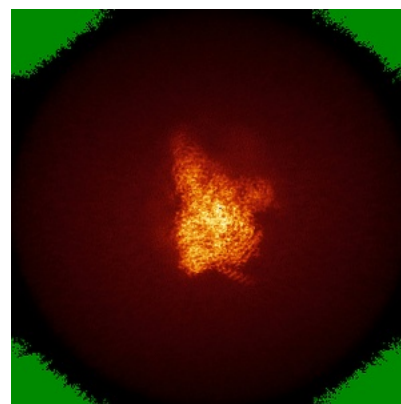
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views ⓘ

6.5.1 Primary map

| X | Y | Z |

The images above show the 3D surface view of the map at the recommended contour level 0.02. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.6 Mask visualisation ⓘ

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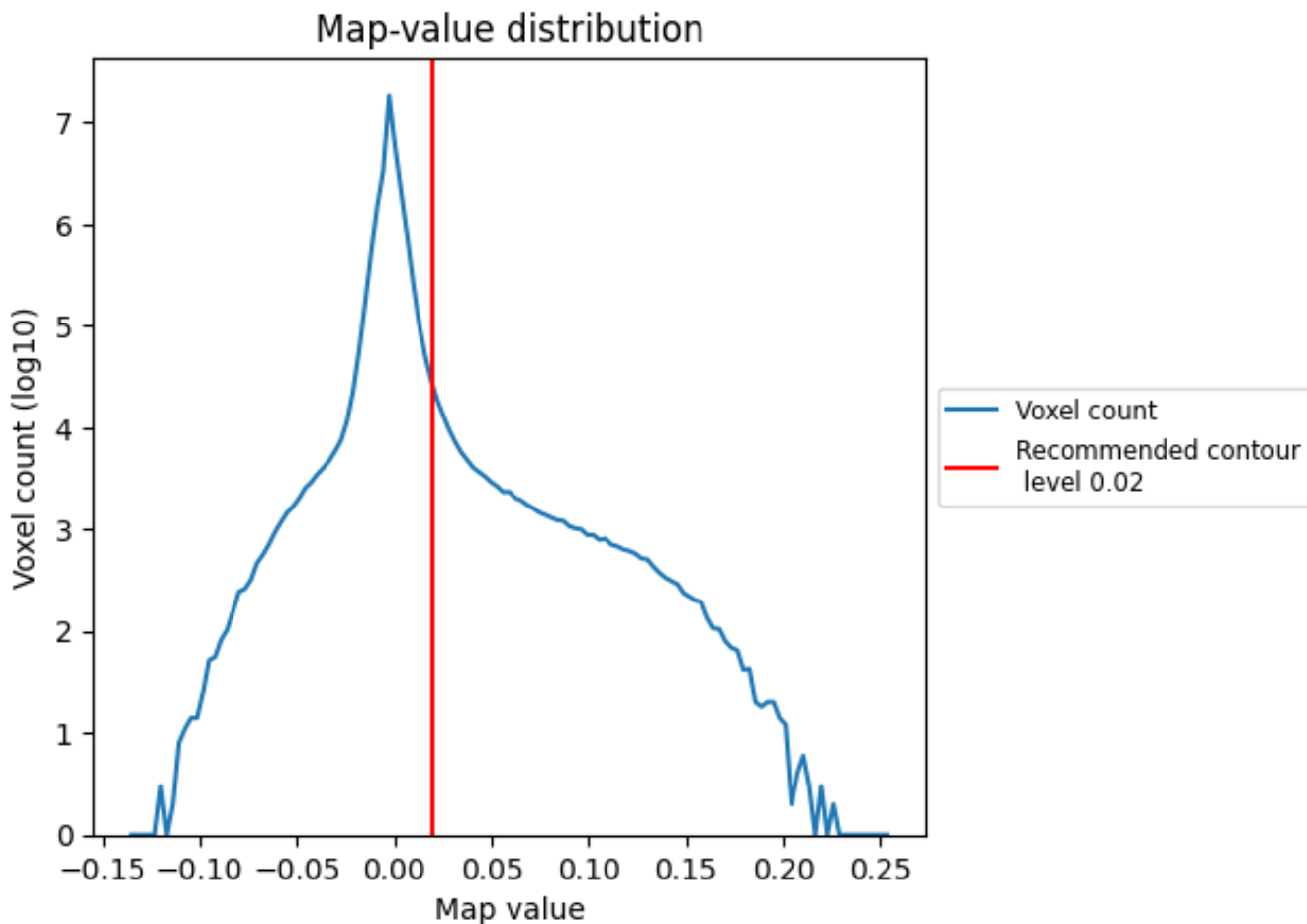
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2 Volume estimate ⓘ
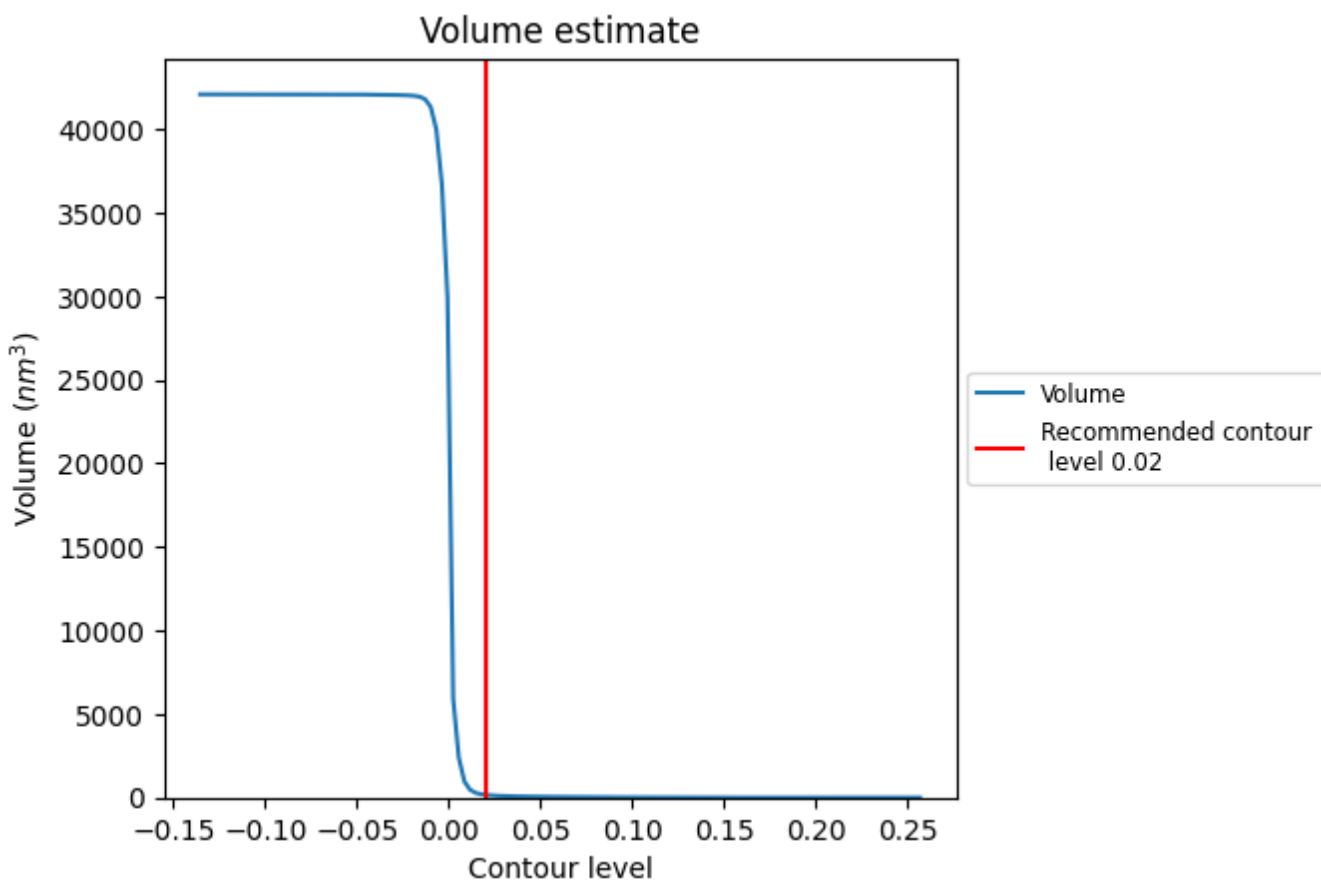
The volume at the recommended contour level is 166 nm³; this corresponds to an approximate mass of 150 kDa.

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 ⓘ
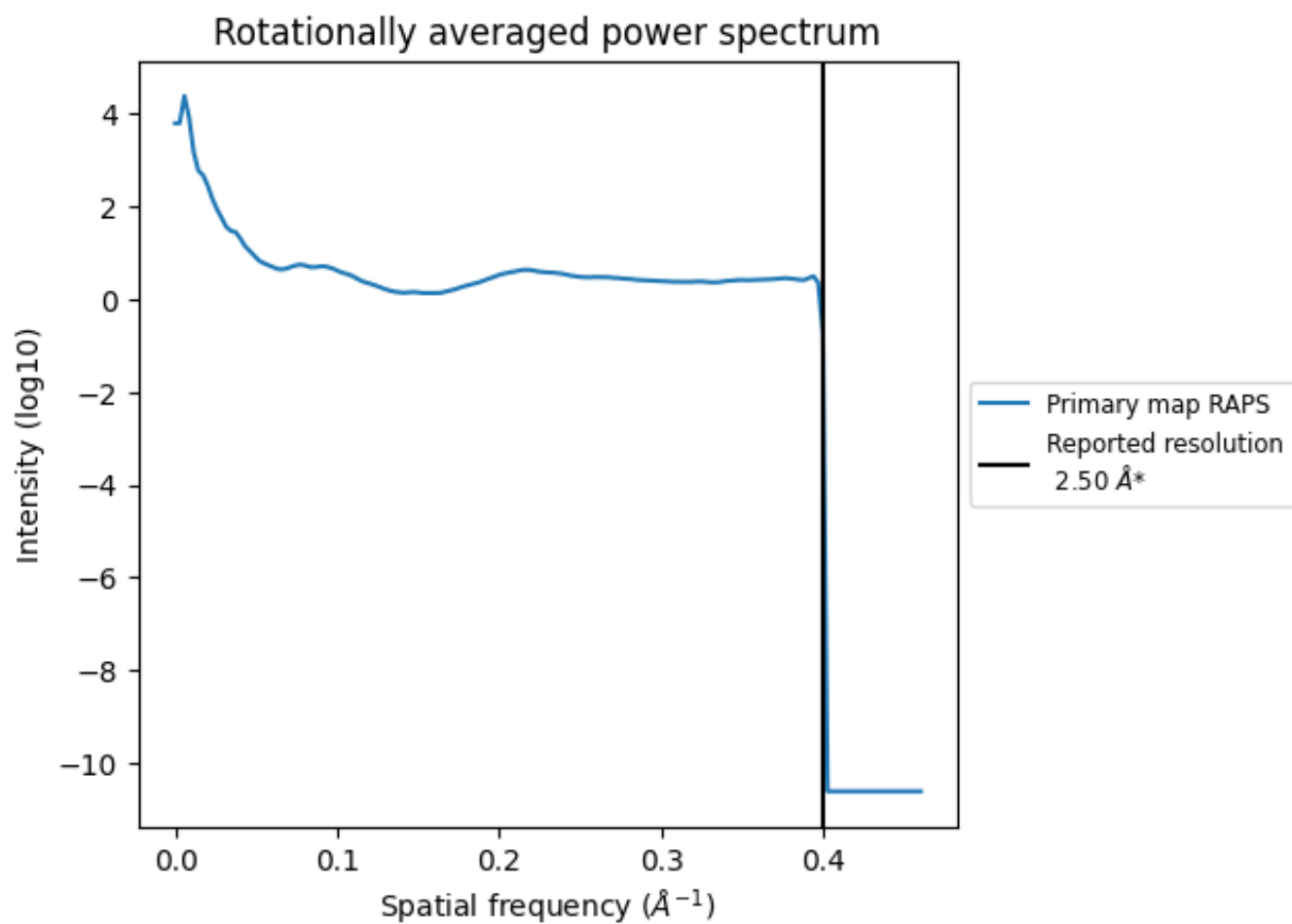
*Reported resolution corresponds to spatial frequency of 0.400 Å^{-1}

8 Fourier-Shell correlation ⓘ

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

9   Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-31334 and PDB model 7EVO. Per-residue inclusion information can be found in section 3 on page 9.

9.1   Map-model overlay ⓘ
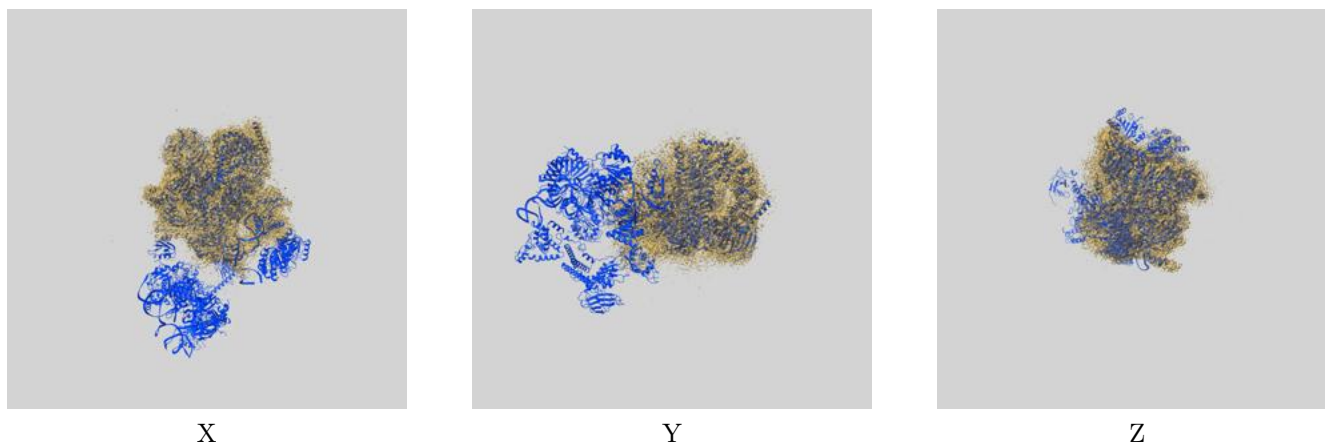
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2   Q-score mapped to coordinate model ⓘ
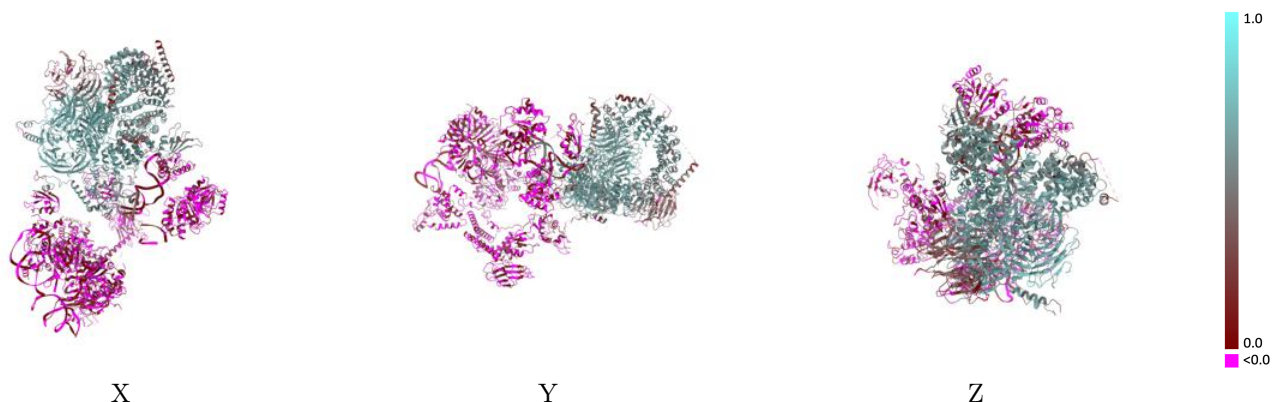
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 ⓘ
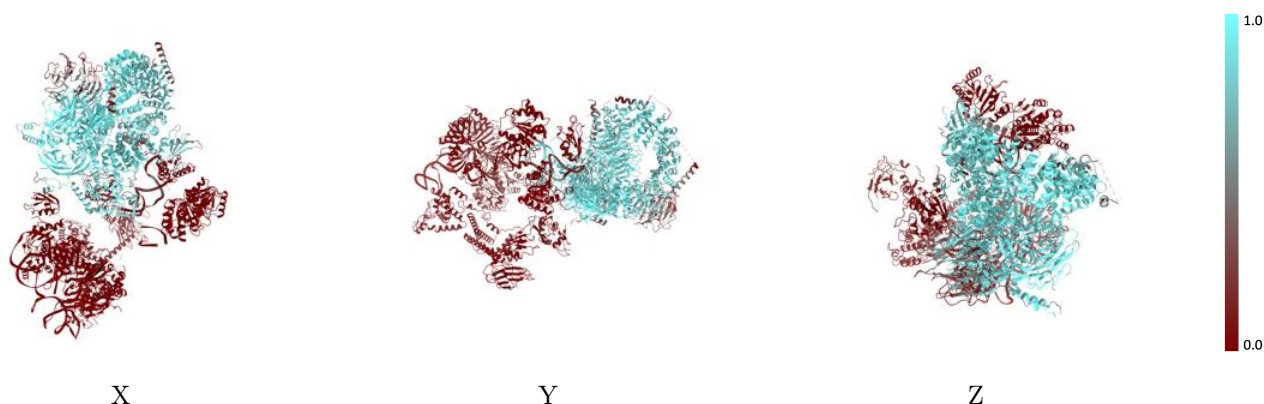
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.02).

9.4 Atom inclusion ⓘ
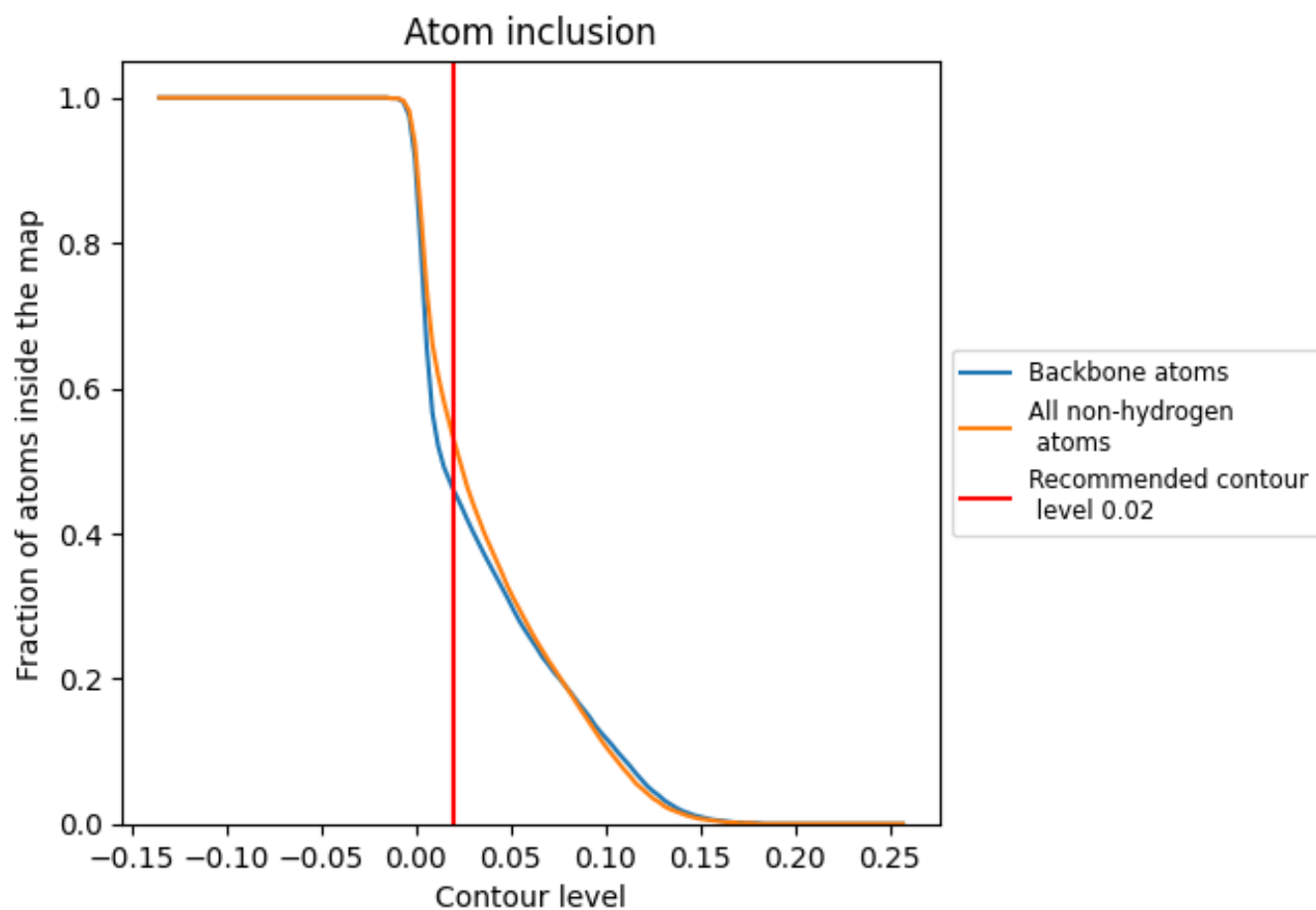
At the recommended contour level, 46% of all backbone atoms, 53% of all non-hydrogen atoms, are inside the map.

9.5   Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5280	0.3670
1	0.8730	0.5840
2	0.6220	0.4210
3	0.8050	0.5730
4	0.0040	0.0160
5	0.9830	0.6830
6	0.9250	0.6450
A	0.0000	0.0030
B	0.0000	-0.0050
C	0.3010	0.2070
D	0.5410	0.3660
E	0.0920	0.0680
F	0.0000	0.0180
G	0.0000	-0.0010
H	0.1450	0.0740
a	0.0000	0.0310
b	0.0000	0.0080
c	0.0000	0.0310
d	0.0000	-0.0240
e	0.0000	0.0410
f	0.0000	0.0290
g	0.0000	0.0440

1.0

0.0
<0.0