



## Full wwPDB EM Validation Report ⓘ

May 13, 2023 – 02:25 pm BST

PDB ID : 7ZNL  
EMDB ID : EMD-14808  
Title : Structure of the human TREX core THO-UAP56 complex  
Authors : Pacheco-Fiallos, F.B.; Vorlaender, M.K.; Plaschka, C.  
Deposited on : 2022-04-21  
Resolution : 3.45 Å (reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

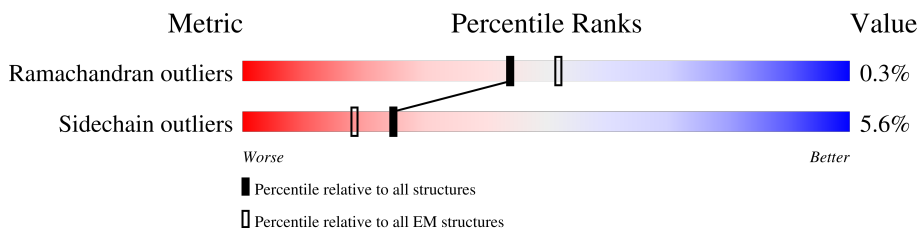
EMDB validation analysis : 0.0.1.dev50  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.32.2

# 1 Overall quality at a glance i

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

The reported resolution of this entry is 3.45 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)
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  $\geq 3$ , 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  $\leq 5\%$ . 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	A	657	
1	I	657	
1	a	657	
1	i	657	
2	B	1593	
2	J	1593	
2	b	1593	
2	j	1593	
3	C	351	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	K	351	88% 77% 11% 12%
3	c	351	88% 77% 11% 12%
3	k	351	88% 77% 11% 12%
4	E	683	78% 78% 21%
4	M	683	80% 80% 20%
4	e	683	77% 77% 23%
4	m	683	80% 80% 20%
5	F	341	99% 98% ..
5	N	341	99% 98% ..
5	f	341	96% 96% ..
5	n	341	99% 98% ..
6	G	204	75% 79% 21%
6	O	204	80% 80% 20%
6	g	204	79% 79% 21%
6	o	204	80% 80% 20%
7	H	428	32% 40% 60%
7	P	428	40% 40% 60%
7	h	428	40% 40% 60%
7	p	428	40% 40% 60%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 84821 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called THO complex subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	356	Total 2889	C 1864	N 481	O 531	S 13	0	0
1	I	363	Total 2935	C 1894	N 489	O 539	S 13	0	0
1	a	356	Total 2889	C 1864	N 481	O 531	S 13	0	0
1	i	363	Total 2936	C 1894	N 489	O 540	S 13	0	0

- Molecule 2 is a protein called THO complex subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	919	Total 7012	C 4486	N 1206	O 1278	S 42	0	0
2	J	914	Total 6956	C 4452	N 1194	O 1268	S 42	0	0
2	b	919	Total 7012	C 4486	N 1206	O 1278	S 42	0	0
2	j	914	Total 6956	C 4452	N 1194	O 1268	S 42	0	0

- Molecule 3 is a protein called THO complex subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	309	Total 2433	C 1538	N 423	O 457	S 15	0	0
3	K	309	Total 2433	C 1538	N 423	O 457	S 15	0	0
3	c	309	Total 2433	C 1538	N 423	O 457	S 15	0	0
3	k	309	Total 2433	C 1538	N 423	O 457	S 15	0	0

- Molecule 4 is a protein called THO complex subunit 5 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	538	Total	C	N	O	S	0	0
			3763	2383	663	696	21		
4	M	549	Total	C	N	O	S	0	0
			4134	2634	725	752	23		
4	e	527	Total	C	N	O	S	0	0
			3674	2327	649	679	19		
4	m	549	Total	C	N	O	S	0	0
			4134	2634	725	752	23		

- Molecule 5 is a protein called THO complex subunit 6 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	337	Total	C	N	O	S	0	0
			2604	1647	459	483	15		
5	N	337	Total	C	N	O	S	0	0
			2604	1647	459	483	15		
5	f	329	Total	C	N	O	S	0	0
			2537	1604	448	470	15		
5	n	337	Total	C	N	O	S	0	0
			2604	1647	459	483	15		

- Molecule 6 is a protein called THO complex subunit 7 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	G	161	Total	C	N	O		0	0
			800	478	161	161			
6	O	164	Total	C	N	O	S	0	0
			1129	696	210	216	7		
6	g	161	Total	C	N	O		0	0
			800	478	161	161			
6	o	164	Total	C	N	O	S	0	0
			1129	696	210	216	7		

- Molecule 7 is a protein called Spliceosome RNA helicase DDX39B.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	170	Total	C	N	O	S	0	0
			1398	888	245	261	4		
7	P	170	Total	C	N	O	S	0	0
			1398	888	245	261	4		
7	h	170	Total	C	N	O	S	0	0
			1398	888	245	261	4		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	p	170	1398	888	245	261	4	0	0



GLU GLN  
ILE ILE  
GLU ASN  
VAL PRO  
PHE LEU  
ALA ASN  
LYS LYS  
LEU LEU  
GLY LEU  
LEU LEU  
TRP GLN

ASN  
LEU  
LEU  
ASN  
ALA  
LEU  
ASN  
LYS  
SER  
SER  
SER  
LEU  
LEU  
ALA  
GLU  
SER  
LEU  
TRP  
LEU  
ALA  
GLU  
SER  
LEU  
LEU  
THR  
TYR  
PRO  
LEU  
LEU  
GLU  
GLN

• Molecule 1: THO complex subunit 1



MET SER  
PRO THR  
THR PRO  
PRO PRO  
PRO LEU  
PHE LEU  
SER SER  
L10  
E70  
E12  
A13  
R44  
R16  
R46  
F17  
T18  
K19  
S20  
T21  
R22  
GLU  
ALA  
LEU  
ASN  
GLN  
ILE  
GLU  
CYS  
SER  
N29  
I30  
K31  
P32  
L33  
L34  
S35  
ARG  
T36  
F37  
S38  
GLN  
VAL  
PRO  
GLY  
SER  
E44  
M45  
E46  
K47  
K48  
C49  
T50  
L51  
D52  
Q53  
A54  
F55  
R56  
G57  
I58  
L59  
E60

E61  
E62  
I63  
I64  
M65  
SER  
SER  
SER  
CYS  
E70  
N71  
V72  
L73  
A74  
I75  
I76  
S77  
L78  
A79  
I80  
S20  
G81  
G82  
V83  
T84  
GLU  
ILE  
CYS  
THR  
S91  
T92  
P93  
F94  
V95  
L96  
L97  
G98  
D99  
V100  
L101  
D102  
C103  
L104  
P105  
L106  
D107  
Q108  
C109  
T110  
T111  
I112  
F113  
T114  
F115  
V116  
E117  
N118  
V120

A121  
T122  
W123  
LYS  
SER  
ASN  
THR  
PHE  
TYR  
SER  
SER  
GLY  
K133  
N134  
Y135  
L136  
L137  
R138  
M139  
C140  
N141  
D142  
L143  
L144  
R145  
R146  
L147  
S148  
K149  
S150  
Q151  
N152  
T153  
V154  
F155  
C156  
G157  
R158  
I159  
Q160  
L161  
F162  
L163  
A164  
R165  
L166  
F167  
PRO  
LEU  
SER  
GLY  
GLU  
LYS  
SER  
SER  
LYS  
SER  
GLY  
ASN  
LEU  
GLN  
GLM

PHE  
ASN  
LEU  
GLU  
ASN  
VAL  
THR  
VAL  
PHE  
ASN  
THR  
ASN  
GLU  
GLN  
SER  
SER  
LEU  
GLY  
GLM  
LYS  
HIS  
THR  
GLU  
ASP  
ARG  
GLU  
GLY  
MET  
ASP  
VAL  
GLU  
GLY  
GLY  
MET  
GLY  
GLY  
MET  
GLY  
ASP  
GLU  
ALA  
PRO  
THR  
THR  
CYS  
S227  
I228  
P229  
L230  
D231  
Y232  
N233  
L234  
Y235  
R236  
K237  
F238  
L239  
S240

L241  
Q242  
D243  
Y244  
F245  
R246  
N247  
P248  
V249  
Q250  
C251  
Y252  
K254  
I255  
S256  
W257  
K258  
T259  
F260  
L261  
K262  
Y263  
S264  
E265  
E266  
V267  
L268  
A269  
V270  
F271  
K272  
S273  
Y274  
K275  
L276  
D277  
D278  
THR  
GLN  
ALA  
SER  
ARG  
LYS  
K285  
M286  
E287  
E288  
L289  
THR  
THR  
GLY  
GLY  
GLU  
HIS  
V296  
Y297  
F298  
A299  
K300

F301  
L302  
T303  
S304  
E305  
K306  
L307  
M308  
D309  
L310  
Q311  
L312  
S313  
D314  
S315  
N316  
F317  
R318  
R319  
H320  
I321  
L322  
L323  
Q324  
Y325  
L326  
I327  
L328  
F329  
Q330  
Y331  
L332  
K333  
G334  
GLN  
VAL  
PHE  
LYS  
SER  
SER  
N342  
Y343  
V344  
L345  
T346  
D347  
E348  
Q349  
S350  
L351  
W352  
I353  
E354  
D355  
T356  
T357  
K358  
S359  
V360

Y361  
Q362  
L363  
L364  
S365  
E366  
N367  
P368  
P369  
D370  
G371  
E372  
R373  
F374  
S375  
K376  
M377  
V378  
E379  
H380  
I381  
L382  
N383  
T384  
E385  
S386  
N387  
W388  
N389  
S390  
W391  
K392  
ASN  
GLU  
GLY  
CYS  
PRO  
SER  
PHE  
VAL  
LYS  
GLU  
ARG  
E458  
H459  
M460  
P461  
T462  
L463  
E464  
E465  
F466  
F467  
E468  
E469  
A470  
L471  
E472  
Q473  
A474  
ASP  
PRO  
GLU  
ASN  
MET  
VAL  
A418  
P419  
E420

D421  
F422  
L423  
G424  
LYS  
GLY  
PRO  
THR  
K429  
R430  
L431  
L432  
M433  
C434  
M435  
E436  
E437  
L438  
T439  
R440  
L441  
M442  
M443  
L444  
C445  
PRO  
ASP  
MET  
SER  
LEU  
VAL  
ALA  
CYS  
GLY  
THR  
THR  
THR  
THR  
THR  
THR  
THR  
GLY  
GLN  
LEU  
ILE  
SER  
MET  
LEU  
VAL  
PHE  
LEU  
SER  
ASP  
SER  
MET  
SER  
LEU  
LEU  
LEU  
LEU  
ALA  
GLY  
SER  
TRP  
GLN  
PRO  
ASP  
GLN  
VAL  
GLY  
VAL  
HIS  
ALA  
THR  
PRO  
GLY  
ASN  
GLU  
ILE  
SER  
MET  
ALA  
VAL  
PHE  
ALA  
ASN  
LYS  
SER  
GLY  
LEU  
LEU  
LEU  
GLN  
TRP  
LEU  
LEU  
VAL  
SER  
SER  
SER  
MET  
GLY  
SER

GLU  
M482  
E483  
Y484  
K485  
A486  
V487  
M488  
M489  
S490  
M491  
G493  
W494  
R495  
A496  
L497  
R498  
L499  
A501  
R502  
R503  
S504  
P505  
H506  
F507  
F508  
Q509  
P510  
T511  
N512  
Q513  
Q514  
F515  
K516  
S517  
L518  
P519  
E520  
Y521  
L522  
E523  
N524  
M525  
V526  
L527  
K528  
LEU  
ALA  
LYS  
GLU  
LEU  
PRO  
PRO  
PRO  
SER  
GLU  
ILE

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

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

• Molecule 1: THO complex subunit 1



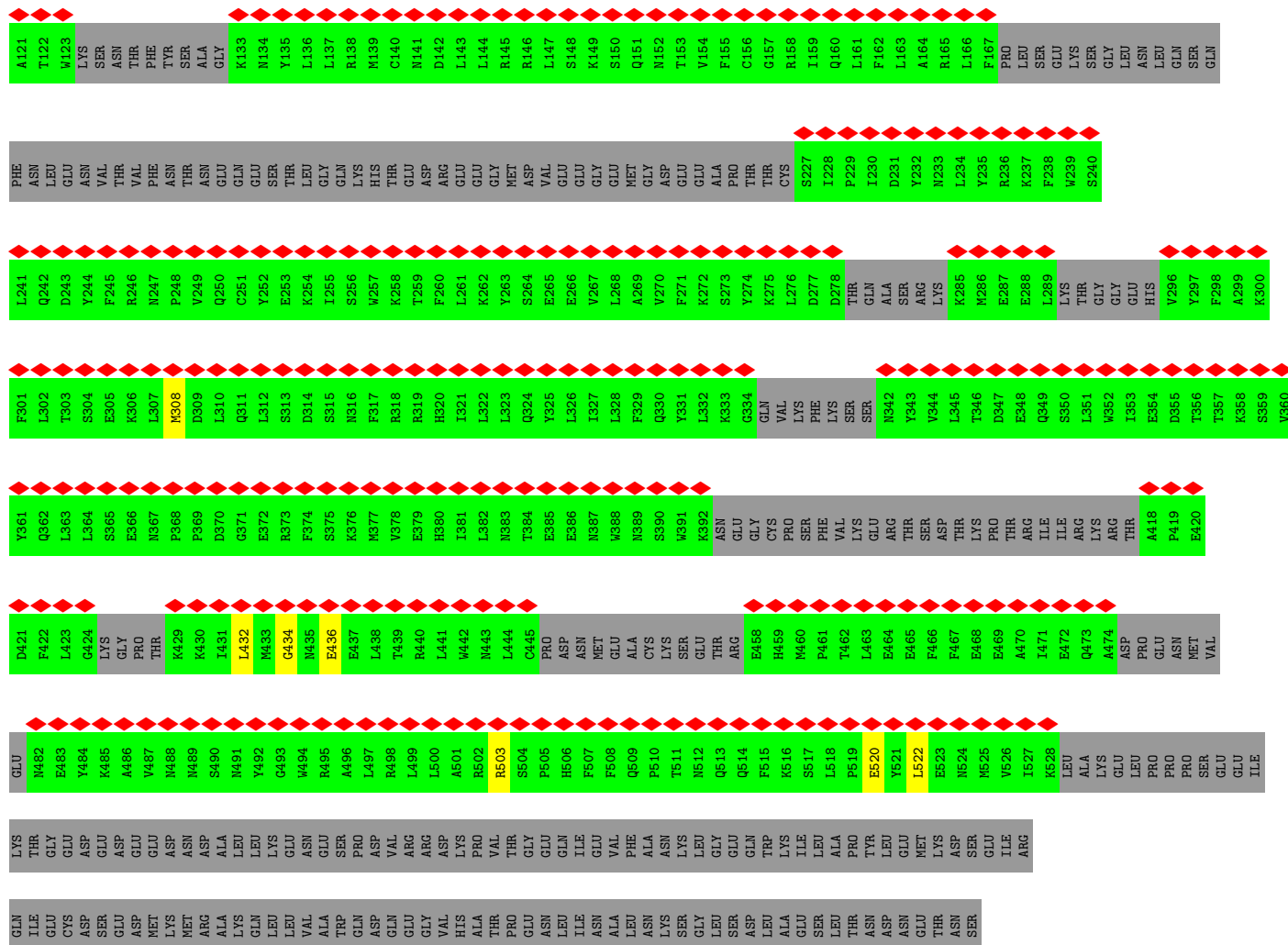


ALA	THR	TRP	LYS	ASN	THR	PHE	THR	GLY	L10	P11	E12	A13	R14	T15	R16	F17	T18	K19	S20	T21	R22	GLU	ALA	LEU	ASN	ASN	LYS	N29	I30	K31	P32	L33	L34	S35	T36	F37	S38	GLN	VAL	PRO	GLY	GLY	SER	E44	N45	E46	K47	K48	C49	T50	L51	D52	Q53	A54	F55	R56	G57	I58	L59	E60							
E61	E62	I63	I64	N65	HIS	SER	SER	CYS	N71	V72	L73	A74	I75	L76	S77	L78	A79	I80	G81	G82	V83	T84	GLU	GLY	ILE	CYS	THR	ALA	S91	T92	P93	F94	V95	L96	L97	G98	D99	V100	L101	D102	C103	L104	P105	L106	D107	Q108	C109	D110	T111	I112	A113	F114	G115	V116	E117	K118	N119	V120									
ALA	THR	TRP	LYS	ASN	THR	PHE	THR	GLY	K133	N134	I135	L136	L137	L138	M139	C140	M141	D142	L143	L144	R145	R146	L147	S148	K149	S150	Q151	N152	T153	V154	F155	C156	G157	R158	I159	Q160	L161	F162	L163	A164	R165	L166	F167	PRO	LEU	SER	GLU	LYS	LYS	THR	GLY	GLY	ASN	LEU	LEU	GLN	SER	GLN	GLN								
PHE	ASN	LEU	GLU	ASN	VAL	THR	VAL	THR	PHE	ASN	THR	ASN	GLN	GLN	C251	Y252	E253	K254	I255	S256	W257	K258	F259	L261	K262	Y263	S264	E265	E266	V267	L268	A269	V270	F271	K272	S273	Y274	K275	L276	D277	I278	THR	GLN	ALA	SER	ARG	ARG	LYS	LYS	MET	Y234	L235	R236	K237	F238	W239	S240										
L241	Q242	D243	Y244	R245	R246	N247	P248	V249	Q250	C251	Y252	E253	K254	I255	S256	W257	K258	F259	L261	K262	Y263	S264	E265	E266	V267	L268	A269	V270	F271	K272	S273	Y274	K275	L276	D277	I278	THR	GLN	ALA	SER	ARG	ARG	LYS	LYS	MET	Y234	L235	R236	K237	F238	W239	S240															
F301	L302	T303	S304	E305	K306	L307	M308	D309	L310	Q311	L312	S313	D314	S315	N316	F317	R318	R319	H320	I321	L322	L323	Q324	Y325	L326	I327	L328	F329	Q330	Y331	L332	K333	G334	GLN	VAL	LYS	PHE	LYS	SER	SER	N342	Y343	V344	L345	T346	D347	E348	Q349	S350	L351	W352	I353	E354	D355	T356	T357	K358	S359	V360								
Y361	Q362	L363	L364	S365	E366	N367	P368	P369	D370	G371	E372	R373	F374	S375	K376	M377	V378	E379	H380	I381	L382	N383	T384	E385	E386	N387	W388	N389	S390	W391	K392	N393	E394	GLY	CYS	PRO	PHE	VAL	VAL	LYS	GLU	ARG	H458	H459	M460	P461	T462	L463	E464	F465	F466	F467	E468	E469	A470	I471	E472	Q473	A474	ASP	PRO	GLU	ASN	VAL	A418	P419	E420
D421	F422	L423	G424	LYS	PRO	THR	K429	K430	I431	L432	M433	G434	M435	E436	E437	L438	T439	R440	L441	W442	M443	L444	C445	PRO	ASP	ASN	MET	GLU	ALA	ALA	CYS	LYS	GLU	THR	ARG	E458	H459	M460	P461	T462	L463	E464	F465	F466	F467	E468	E469	A470	I471	E472	Q473	A474	ASP	PRO	GLU	ASN	VAL	A418	P419	E420							
GLU	N462	E483	Y484	K485	A486	W487	M488	M489	S490	M491	Y492	G493	W494	R495	A496	L497	R498	L499	L500	A501	R502	E503	S504	P505	H506	F507	F508	Q509	P510	T511	N512	Q513	Q514	F515	K516	S517	L518	P519	E520	Y521	L522	E523	M524	M525	V526	I527	K528	LEU	LEU	ALA	LYS	GLU	LEU	PRO	PRO	PRO	PRO	SER	GLU	GLU	ILE						
LYS	THR	GLY	ASP	GLU	ASP	GLU	LYS	ASP	ASN	ASP	ALA	LYS	GLN	LEU	VAL	ASN	TRP	GLN	ASN	ASP	VAL	THR	GLY	GLU	ASN	GLN	ILE	ILE	ALA	VAL	PHE	ALA	ASN	P510	T511	N512	Q513	Q514	F515	K516	S517	L518	P519	E520	Y521	L522	E523	M524	M525	V526	I527	K528	LEU	LEU	ALA	LYS	GLU	LEU	PRO	PRO	PRO	PRO	SER	GLU	GLU	ILE	
GLN	ILE	GLU	CYS	ASP	GLU	ASP	LYS	MET	MET	ARG	ALA	LYS	GLN	LEU	VAL	ALA	TRP	GLN	ASN	ASP	GLN	THR	GLY	GLU	ASN	LEU	LEU	ALA	VAL	PHE	ALA	ASN	P510	T511	N512	Q513	Q514	F515	K516	S517	L518	P519	E520	Y521	L522	E523	M524	M525	V526	I527	K528	LEU	LEU	ALA	LYS	GLU	LEU	PRO	PRO	PRO	PRO	SER	GLU	GLU	ILE		

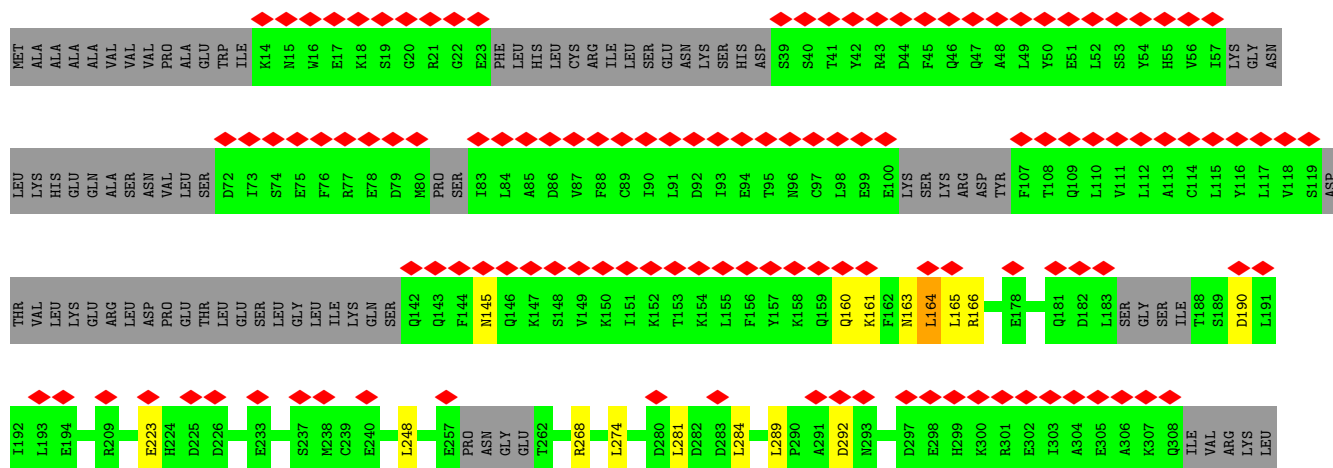
Molecule 1: THO complex subunit 1

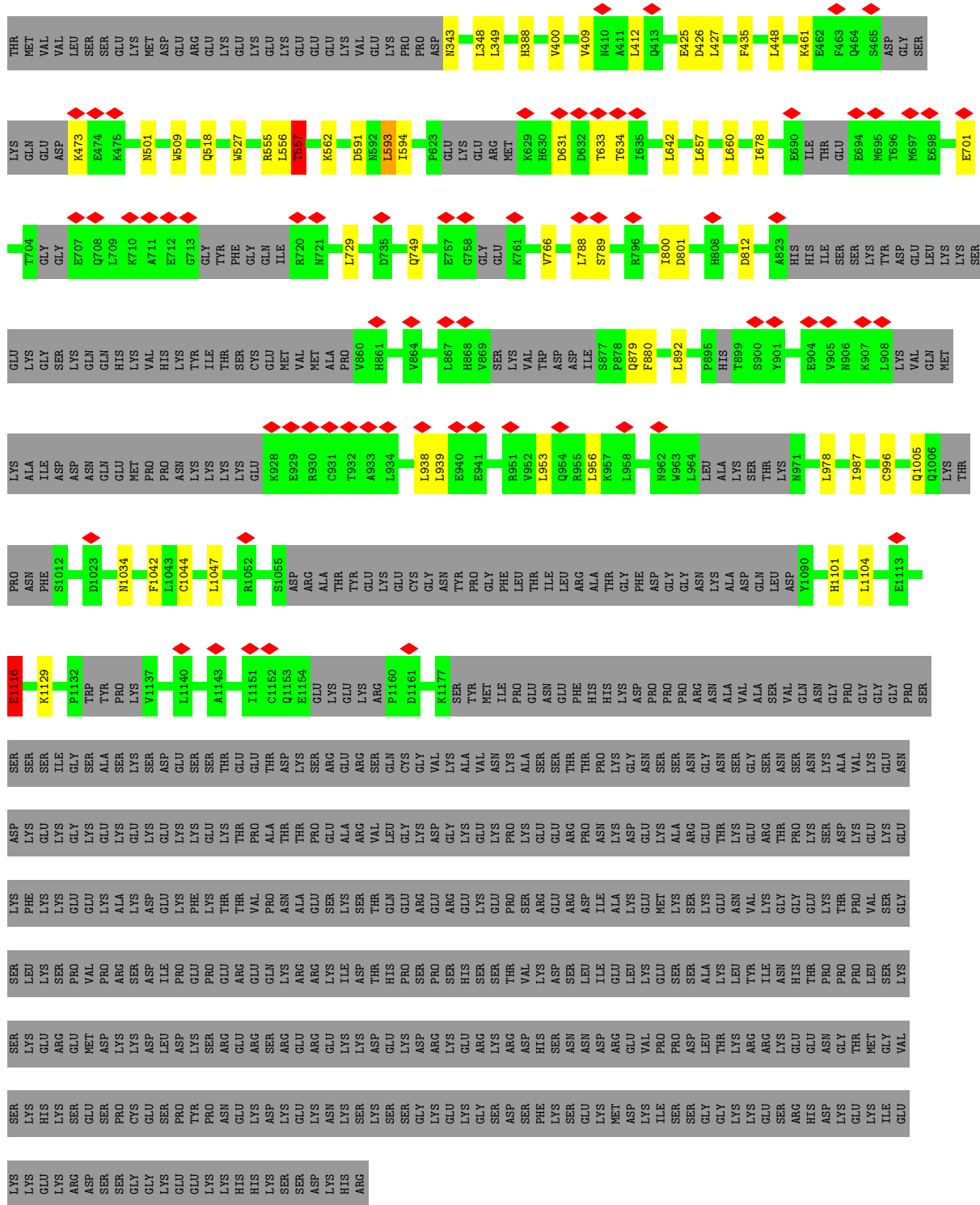


MET	SER	PRO	THR	PRO	PRO	PRO	LEU	PHE	SER	L10	P11	E12	A13	R14	T15	R16	F17	T18	K19	S20	T21	R22	GLU	ALA	LEU	ASN	ASN	LYS	N29	I30	K31	P32	L33	L34	S35	T36	F37	S38	GLN	VAL	PRO	GLY	GLY	SER	E44	N45	E46	K47	K48	C49	T50	L51	D52	Q53	A54	F55	R56	G57	I58	L59	E60
E61	E62	I63	I64	N65	HIS	SER	SER	CYS	E70	N71	V72	L73	A74	I75	L76	S77	L78	A79	I80	G81	G82	V83	T84	GLU	GLY	ILE	CYS	THR	ALA	S91	T92	P93	F94	V95	L96	L97	G98	D99	V100	L101	D102	C103	L104	P105	L106	D107	Q108	C109	D110	T111	I112	A113	F114	G115	V116	E117	K118	N119	V120		

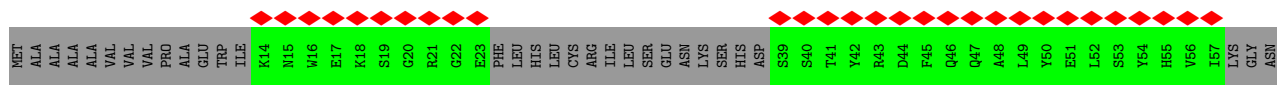
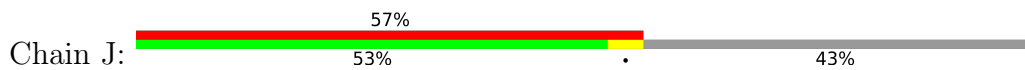


● Molecule 2: THO complex subunit 2





● Molecule 2: THO complex subunit 2





R903	E904	V905	N906	K907	L908	LYS	VAL	GLN	MET	LYS	ALA	ILE	ASP	ASN	GLN	GLU	MET	PRO	PRO	GLU	K928	E929	R930	C931	T932	A933	L934	Q935	D936	K937	L938	L939	E940	E941	E942	K943	K944	Q945	M946	E947	H948	V949	Q950	R951	V952	L953	Q954	R955	L956	K957	L958	E959	K960						
D961	N962	W963	L964	LEU	ALA	LYS	THR	LYS	M971	E972	T973	T974	T975	K976	F977	L978	Q979	L980	C981	F982	P983	P984	R985	C986	I987	F988	S989	A990	I991	D992	A993	V994	Y995	C996	A997	R998	F999	V1000	E1001	L1002	V1003	H1004	Q1005	Q1006	LYS	THR	PRO	ASN	PHE	S1012	T1013	L1014	L1015	C1016	Y1017	D1018	R1019	V1020	
F1021	S1022	S1023	I1024	T1025	Y1026	T1027	V1028	A1029	S1030	C1031	T1032	E1033	M1034	E1035	A1036	S1037	R1038	Y1039	G1040	F1041	R1042	L1043	L1044	C1045	M1046	L1047	E1048	T1049	V1050	T1051	D992	A993	V994	Y995	C996	A997	R998	F999	V1000	E1001	L1002	V1003	H1004	Q1005	Q1006	LYS	THR	PRO	ASN	PHE	S1012	T1013	L1014	L1015	C1016	Y1017	D1018	R1019	V1020
GLY	GLY	ASN	ASN	ALA	ASP	GLN	LEU	ASP	Y1090	E1091	M1092	F1093	R1094	H1095	E1096	V1097	H1098	K1099	W1100	H1101	Y1102	K1103	L1104	T1105	K1106	A1107	S1108	V1109	H1110	C1111	L1112	E1113	T1114	G1115	H1116	Y1117	T1118	H1119	I1120	A1121	M1122	L1123	C1124	H1125	V1126	L1127	T1128	K1129	I1130	L1131	P1132	TRP	TYR	PRO	LYS	V1137	L1138	M1139	L1140
G1141	Q1142	A1143	L1144	E1145	R1146	R1147	V1148	H1149	K1150	I1151	C1152	Q1153	E1154	GLU	LYS	GLU	GLU	LYS	LYS	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	ALA	VAL	VAL									
GLN	ASN	GLY	PRO	GLY	GLY	PRO	GLY	PRO	ASN	LYS	ASP	GLY	ASP	GLY	ASP	GLY	GLU	GLU	LYS	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	VAL	VAL										
ASN	SER	ASN	ASN	ALA	VAL	LYS	GLU	ASN	LYS	ASP	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	VAL	VAL										
THR	PRO	LYS	SER	ASP	THR	LYS	GLY	LYS	PHE	LYS	LYS	GLY	GLY	GLY	GLY	GLY	GLY	PHE	LYS	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	VAL	VAL										
GLY	GLY	GLU	THR	THR	VAL	VAL	GLY	GLY	GLY	LEU	VAL	VAL	VAL	VAL	VAL	ILE	PRO	GLU	THR	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	VAL	VAL										
ASN	HIS	THR	PRO	PRO	PRO	LEU	SER	SER	LYS	GLY	GLU	GLU	MET	LYS	ASP	GLY	ASP	LEU	ASP	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	VAL	VAL										
LYS	GLU	GLU	THR	THR	GLY	VAL	VAL	SER	HIS	LYS	HIS	SER	GLU	GLU	GLU	GLY	GLY	GLU	PRO	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	VAL	VAL										
SER	ARG	HIS	ASP	LYS	GLY	ILE	VAL	VAL	LYS	LYS	GLY	ASP	ASP	GLY	GLY	GLY	GLU	GLU	PRO	F1160	D1161	L1162	Y1163	A1164	L1165	A1166	M1167	G1168	Y1169	S1170	G1171	Q1172	L1173	E1174	K1175	S1176	R1177	GLY	THR	THR	ASN	PRO	PRO	ARG	ALA	VAL	ALA	VAL	VAL										

• Molecule 2: THO complex subunit 2



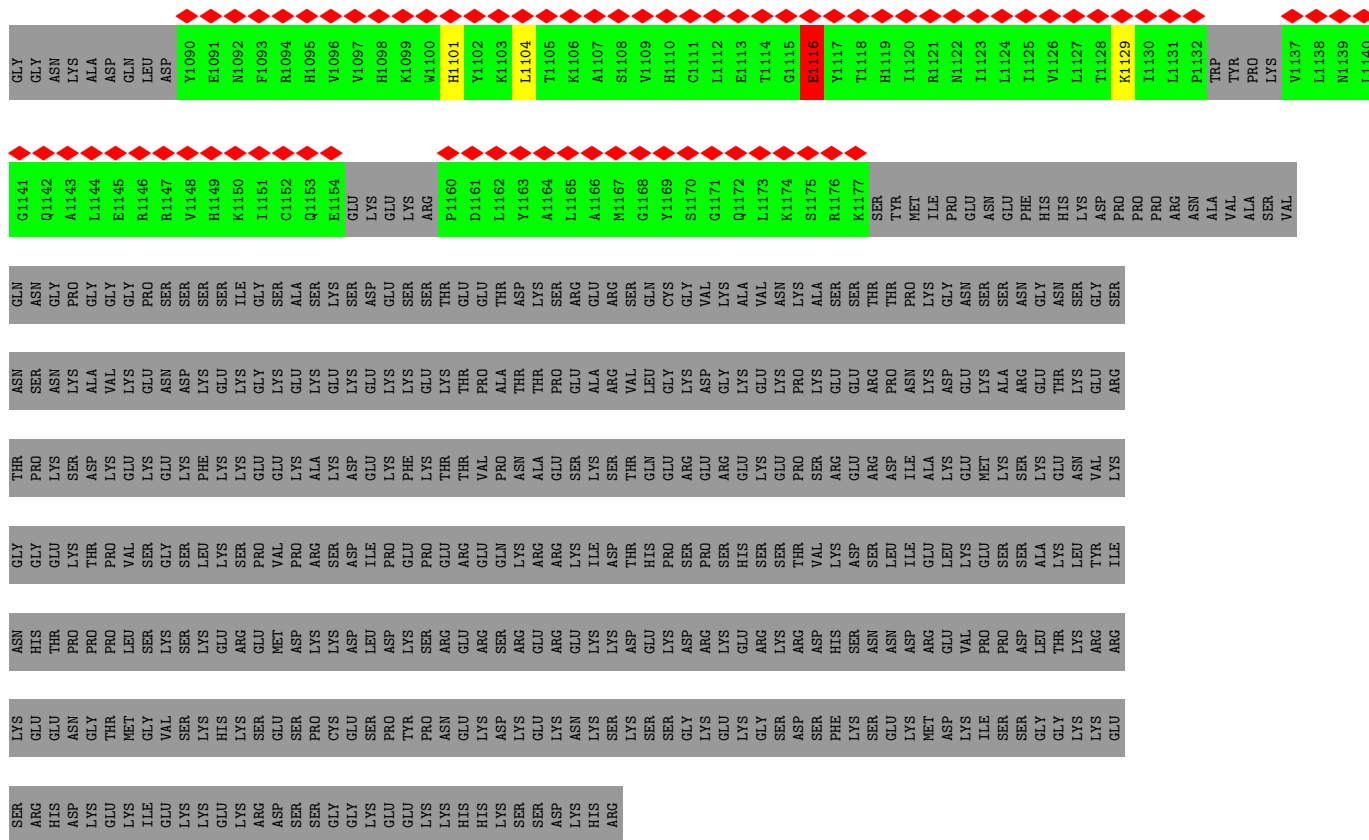
MET	ALA	ALA	ALA	ALA	VAL	VAL	VAL	VAL	ALA	ALA	TRP	ILE	K14	H15	H16	E17	K18	S19	R21	G22	E23	PHE	LEU	HIS	LEU	LEU	CYS	ARG	ILE	LEU	SER	GLU	ASN	LYS	SER	HIS	S39	S40	T41	Y42	D43	D44	F45	Q46	Q47	A48	L49	Y50	E51	L52	S53	Y54	H55	V56	L57	LYS	GLY	ASN
LEU	LYS	HIS	GLU	GLN	ALA	ALA	ASN	VAL	VAL	SER	D72	I73	S74	E75	F76	R77	E78	H80	PRO	SER	I83	L84	A85	D86	W87	F88	C89	I90	L91	D92	I93	E94	T95	N96	C97	L98	E99	E100	L101	SER	LYS	ARG	ASP	TYR	F107	T108	Q109	L110	V111	L112	A113	C114	L115	L116	L117	V118	S119	ASP
THR	VAL	LEU	LYS	GLU	GLU	ARG	LEU	LEU	ASP	PRO	Q142	Q143	F144	M145	Q146	K147	S148	V149	K150	I151	I152	T153	K154	L155	F156	Y157	K158	Q159	Q160	L161	F162	M163	L164	L165	R166	E167	E168	M169	E170	G171	Y172	A173	K174	L175	L176	A177	E178	L179	G180									



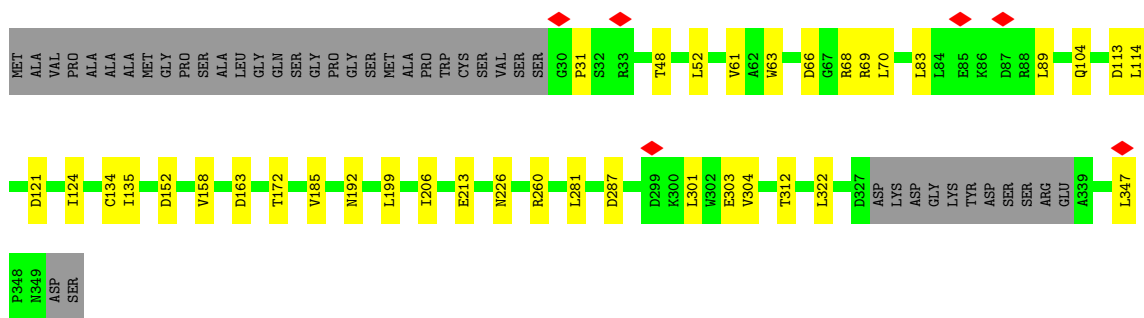
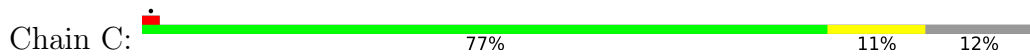


R301	E302	I303	A304	E305	A306	K307	Q308	ILE	VAL	ARG	LEU	THR	MET	VAL	VAL	LEU	SER	SER	GLU	GLY	LYS	MET	ASP	GLU	ARG	GLY	GLY	LYS	PRO	PRO	PRO	ASP	N343	Q344	K345	L346	G347	L348	L349	E350	A351	L352	L353	K354	S355	G356	D357	W358	Q359	H360											
A361	Q362	N363	I364	M365	D366	Q367	M368	P369	P370	Y371	Y372	A373	A374	S375	H376	K377	L378	I379	A380	L381	A382	I383	C384	K385	L386	I387	H388	I389	T390	I391	E392	P393	L394	Y395	R396	R397	L398	G399	V400	P401	K402	G403	A404	K405	G406	S407	P408	V409	N410	A411	L412	Q413	N414	K415	R416	A417	P418	K419	Q420		
A421	E422	F423	F424	E425	D426	L427	R428	R429	D430	V431	F432	N433	M434	F435	C436	L437	L438	G439	P440	H441	L442	S443	H444	D445	P446	I447	L448	F449	A450	K451	V452	V453	V454	L455	G456	K457	S458	F459	M460	K461	F462	F463	Q464	S465	ASP	GLY	SER	SER	LYS	GLN	GLU	ASP	K473	E474	K475	T476	E477	V478	I479	L480	
S481	C482	L483	L484	S485	I486	T487	D488	Q489	D490	L491	L492	P493	S494	L495	S496	L497	M498	D499	C500	N501	A502	C503	M504	S505	E506	E507	L508	W509	G510	M511	F512	K513	T514	F515	P516	Y517	Q518	H519	R520	Y521	R522	L523	Y524	G525	Q526	W527	K528	N529	E530	T531	Y532	N533	S534	H535	P536	L537	V538	V539	K540		
V541	K542	A543	Q544	T545	I546	D547	R548	A549	K550	Y551	I552	M553	K554	R555	L556	T557	K558	E559	N560	V561	K562	P563	S564	G565	R566	Q567	I568	G569	K570	L571	S572	H573	S574	N575	P576	T577	I578	L579	F580	D581	R582	I583	L584	S585	Q586	I587	Q588	K589	N590	D591	N592	L593	I594	T595	P596	V597	L598	D599	S600		
L601	K602	Y603	L604	T605	S606	L607	N608	Y609	D610	V611	L612	A613	Y614	C615	I616	I617	E618	A619	L620	A621	N622	P623	GLU	GLY	ARG	MET	K629	H630	D631	D632	T633	T634	I635	S636	S637	W638	L639	Q640	S641	L642	A643	S644	F645	C646	G647	A648	V649	F650	R651	K652	Y653	P654	I655	D656	L657	A658	G659	L660			
L661	Q662	Y663	V664	A665	N666	Q667	L668	K669	A670	G671	K672	S673	F674	D675	L676	L677	L678	L679	K680	E681	V682	V683	Q684	K685	M686	A687	G688	I689	E690	THR	T693	E694	M695	T696	M697	E698	Q699	L700	E701	A702	M703	S704	GLY	GLY	E707	Q708	L709	K710	A711	E712	GLY	GLY	T719	PHE	GLY	GLN	I720				
M721	T722	K723	K724	S725	S726	Q727	R728	L729	K730	D731	A732	L733	L734	D735	H736	D737	L738	A739	L740	P741	L742	C743	L744	L745	M746	A747	Q748	Q749	R750	M751	G752	L753	L754	F755	Q756	E757	E698	Q699	L700	E701	A702	M703	S704	GLY	GLY	E707	Q708	L709	K710	A711	E712	GLY	GLY	T719	PHE	GLY	GLN	I720	F780		
G781	G782	F783	L784	A785	S786	M787	L788	S789	T790	E791	D792	Y793	L794	K795	R796	V797	L798	A799	I800	D801	L802	L803	C804	N805	E806	F807	H808	T809	P810	H811	D812	A813	L814	F815	F816	L817	S818	R819	P820	M821	A823	HIS	ILE	ILE	SER	SER	LYS	TYR	ASP	GLU	LEU	LYS	LYS	SER	GLU	LYS	LYS	GLY	GLY	SER	
LYS	GLN	HIS	LYS	VAL	HIS	LYS	TYR	THR	ILE	THR	SER	LYS	ALA	GLU	VAL	PRO	H860	H861	E862	A863	V864	V865	S866	L867	H868	SER	LYS	VAL	TRP	ASP	ILE	S877	P878	K879	F880	F881	A882	T883	F884	W885	S886	L887	T888	M889	Y890	D891	L892	A893	V894	P895	HIS	T899	S900	Y901	E902						
R903	E904	Y905	N906	K907	L908	LYS	VAL	GLN	MET	LYS	ALA	ILE	ASP	ASP	GLN	GLU	MET	PRO	Q979	L980	C981	I982	F983	P984	R985	C986	I987	F988	S989	A990	L991	D992	A993	Y994	V995	C996	E997	R998	F999	V1000	E1001	L1002	V1003	H1004	Q1005	Q1006	LYS	THR	PHE	ASN	L953	Q954	R955	L1013	L1014	L1015	L1016	Y1017	D1018	R1019	V1020
D961	N962	W963	L964	LEU	ALA	LYS	SER	THR	THR	LYS	N971	E972	T973	I974	T975	K976	F977	L978	Q979	L980	C981	I982	F983	P984	R985	C986	I987	F988	S989	A990	L991	D992	A993	Y994	V995	C996	E997	R998	F999	V1000	E1001	L1002	V1003	H1004	Q1005	Q1006	LYS	THR	PHE	ASN	L953	Q954	R955	L1013	L1014	L1015	L1016	Y1017	D1018	R1019	V1020
F1021	S1022	D1023	I1024	I1025	Y1026	T1027	V1028	A1029	S1030	C1031	T1032	E1033	M1034	E1035	A1036	S1037	R1038	Y1039	G1040	R1041	F1042	L1043	C1044	M1045	M1046	L1047	E1048	T1049	V1050	T1051	R1052	W1053	W1054	S1055	ASP	ARG	ALA	THR	TYR	GLU	LYS	CYS	GLY	ASN	TYR	PRO	GLY	PHE	THR	LEU	THR	ILE	ARG	ALA	THR	GLY	PHE	ASP			

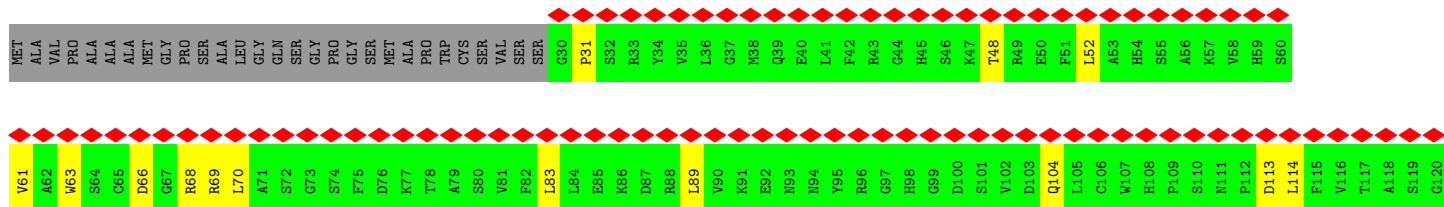
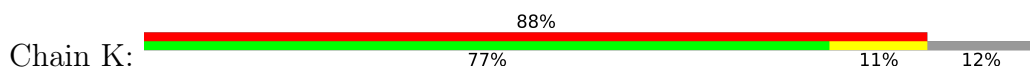


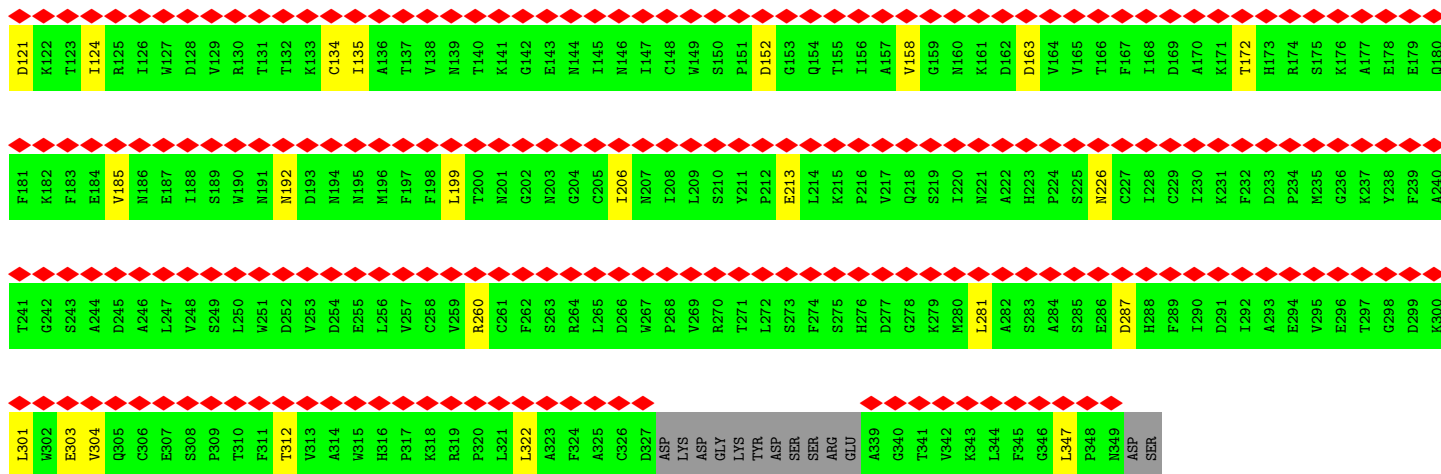


• Molecule 3: THO complex subunit 3

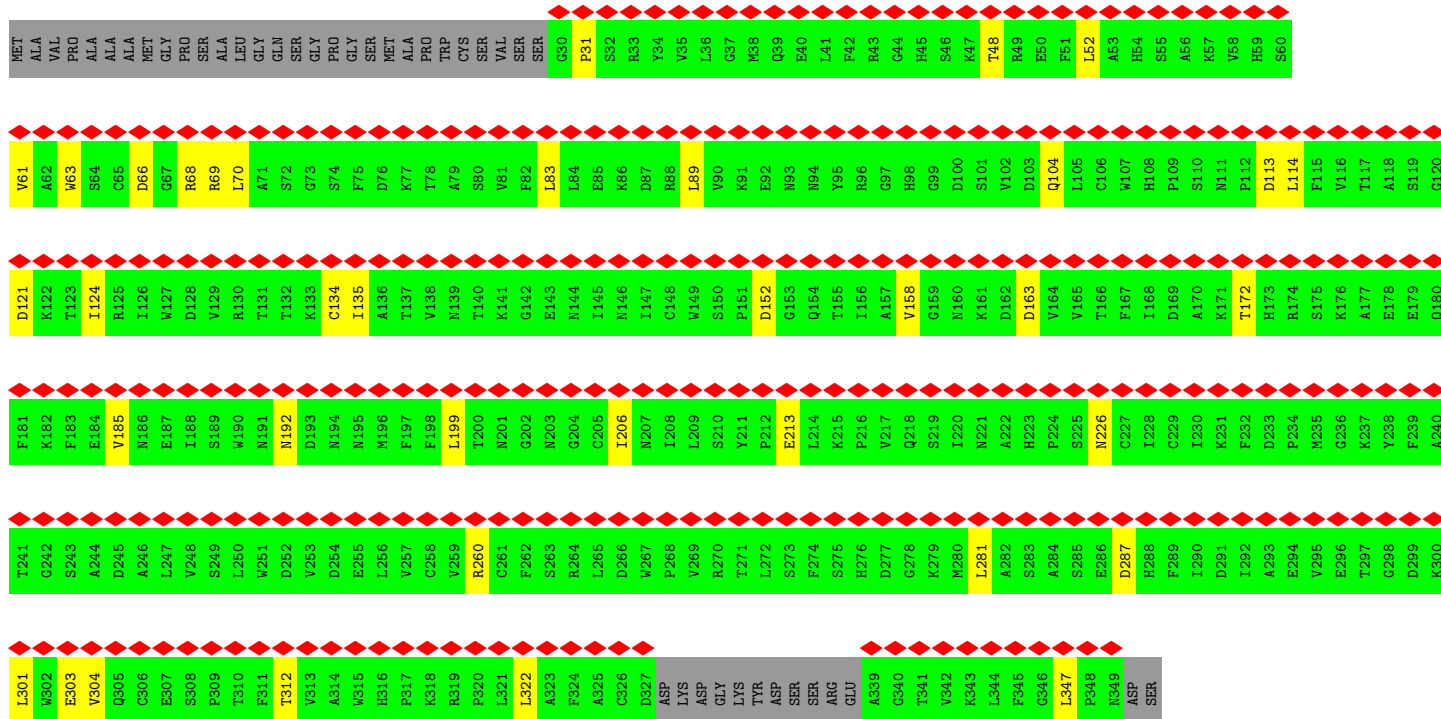
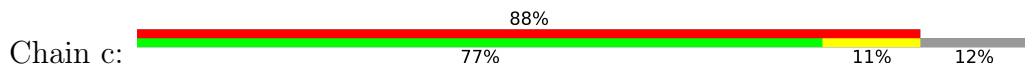


• Molecule 3: THO complex subunit 3

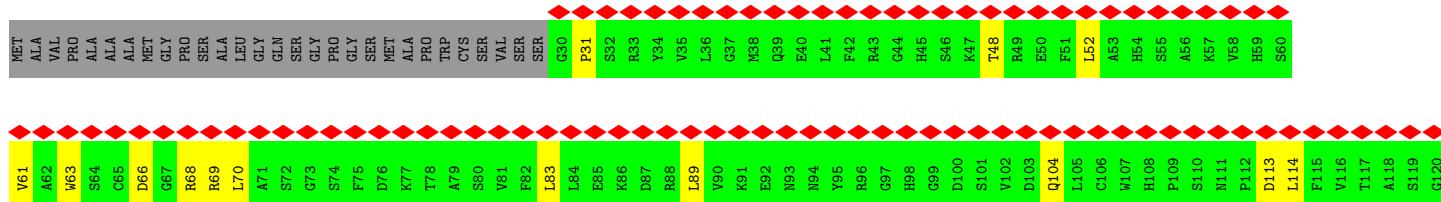
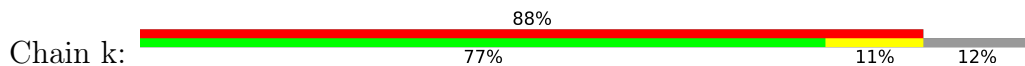




• Molecule 3: THO complex subunit 3

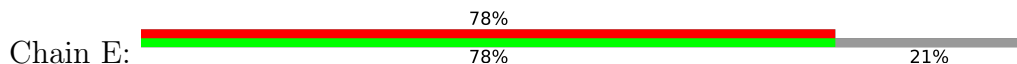


• Molecule 3: THO complex subunit 3

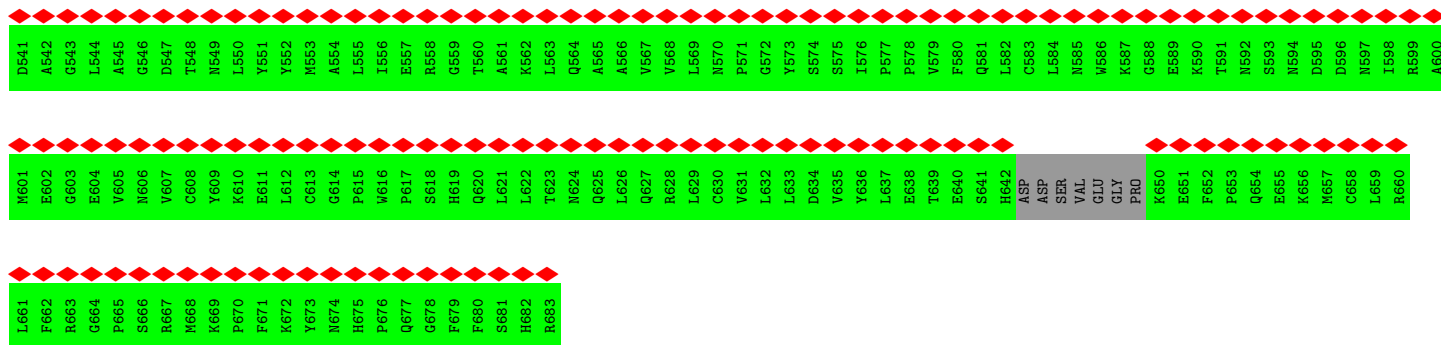


T481	R482	V483	Q484	S485	R486	L487	A488	H489	K491	Q492	F493	A494	S495	L496	E497	H498	G499	I500	V501	P502	V503	T504	S505	D506	C507	Q508	Y509	L510	F511	P512	A513	K514	V515	V516	S517	R518	L519	V520	K521	W522	V523	T524	V525	A526	H527	E528	D529	V530	M531	E532	L533	H534	F535	T536	L479	K480	T481	R482	V483	Q484	S485	R486	L487	A488	H489	K491	Q492	F493	A494	S495	L496	E497	H498	G499	I500	V501	P502	V503	T504	S505	D506	C507	Q508	Y509	L510	F511	P512	A513	K514	V515	V516	S517	R518	L519	V520	K521	W522	V523	T524	V525	A526	H527	E528	D529	V530	M531	E532	L533	H534	F535	T536	L479	K480
F181	K182	F183	E184	V185	N186	E187	I188	S189	W190	M191	N192	D193	N194	N195	M196	F197	F198	L199	T200	N201	G202	N203	G204	C205	T206	N207	I208	L209	S210	Y211	P212	E213	Q154	T155	I156	V217	V158	G159	N160	K161	D162	D163	V164	V165	V166	F167	I168	D169	A170	K171	T172	H173	R174	S175	K176	A177	E178	E179	A240																																																						
T241	G242	S243	A244	D245	A246	L247	V248	S249	L250	W251	D252	V253	D254	E255	L256	V257	C258	V259	R260	C261	F262	S263	R264	L265	D266	W267	P268	V269	R270	T271	L272	S273	F274	S275	H276	D277	G278	K279	M280	L281	S282	S283	A284	S285	E286	D287	H288	F289	I290	D291	I292	A293	E294	V295	E296	T297	G298	D299	K300																																																						
L301	W302	E303	V304	Q305	C306	E307	S308	P309	T310	F311	T312	V313	A314	W315	H316	P317	K318	R319	P320	L321	L322	A323	F324	A325	C326	D327	ASP	LYS	GLY	LYS	TYR	E40	S38	E39	SER	E42	V43	D44	L45	R46	D47	PRO	GLY	K343	L344	F345	G346	L347	P348	N349	ASP	SER	Q60																																																												

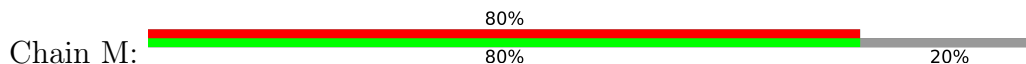
• Molecule 4: THO complex subunit 5 homolog



MET	SER	GLU	SER	LYS	ARG	LYS	LYS	PRO	LYS	VAL	ILE	ARG	ASP	GLY	ALA	PRO	ALA	GLY	LYS	ARG	LYS	ASP	ASN	ARG	ARG	ASP	THR	GLU	GLN	GLY	GLY	TYR	S38	E39	SER	E42	V43	D44	L45	R46	D47	PRO	GLY	K343	L344	F345	G346	L347	P348	N349	ASP	SER	Q60																																																												
E61	L62	Q63	H64	L65	M66	A67	E68	I69	Q70	D71	L72	K73	S74	R75	GLY	GLY	LYS	ASP	W80	A81	I82	E83	I84	E85	E86	H87	R88	I89	Q90	S91	C92	V93	H94	F95	M96	T97	L98	K99	K100	L101	N102	R103	L104	A105	H106	I107	S108	L109	K110	G111	G112	R113	D114	Q115	T116	H117	E118	A119	K120																																																						
Q121	K122	V123	D124	A125	V126	H127	L128	Q129	L130	Q131	M132	L133	L134	V135	L136	V137	M138	H139	L140	A202	R203	K204	L205	R206	E207	C208	L209	S210	N211	K212	E213	K214	L215	L216	K217	E218	L219	E220	V221	K222	K223	E224	V225	L226	S227	S228	L229	Q230	P231	R232	L233	M234	S235	L236	M237	Q238	A239	S240																																																							
L241	P242	V243	Q244	E245	Y246	L247	PHE	PRO	PRO	PHE	D252	Q253	A254	H255	K256	Q257	Y258	E259	T260	A261	R262	H263	L264	P265	P266	P267	L268	Y269	V270	L271	F272	V273	Q274	A275	T276	A277	Y278	G279	Q280	A281	C282	D283	K284	T285	L286	S287	V288	A289	I290	E291	G292	S293	V294	D295	E296	A297	K298	A299	LEU																																																						
PHE	LYS	PRO	PRO	ASP	GLN	THR	A189	ASP	ASP	ASP	ASP	ALA	GLU	GLU	GLN	THR	THR	ARG	ARG	ARG	PRO	THR	LEU	VAL	L333	D334	D335	K336	R337	K338	E339	H340	L341	L342	L343	R344	H344	P345	L346	S347	V348	M349	L350	D351	L352	K353	C354	K355	D356	D357	S358	V359	L360																																																												
H361	L362	T363	F364	Y365	Y366	L367	M368	N369	L370	N371	I372	M373	T374	V375	K376	A377	K378	V379	T380	T381	A382	M383	E384	L385	I386	T387	P388	I389	S390	A391	G392	D393	L394	L395	S396	P397	D398	S399	V400	L401	S402	C403	L404	Y405	P406	G407	D408	H409	Q410	K411	K412	T413	P414	M415	P416	A417	M418	Q419	Y420																																																						
Q421	F422	D423	K424	V425	GLY	ILE	LEU	THR	L430	S431	D432	Y433	V434	L435	E436	L437	G438	H439	P440	Y441	W442	W443	W444	Q445	K446	L447	G448	G449	L450	H451	F452	P453	LYS	GLU	GLN	PRO	GLN	THR	VAL	I462	A463	D464	H465	S466	L467	S468	A469	S470	H471	M472	E473	T474	T475	M476	L477	L478	L479	K480																																																							
T481	R482	V483	Q484	S485	R486	L487	A488	H489	K491	Q492	F493	A494	S495	L496	E497	H498	G499	I500	V501	P502	V503	T504	S505	D506	C507	Q508	Y509	L510	F511	P512	A513	K514	V515	V516	S517	R518	L519	V520	K521	W522	V523	T524	V525	A526	H527	E528	D529	V530	M531	E532	L533	H534	F535	T536	L479	K480	T481	R482	V483	Q484	S485	R486	L487	A488	H489	K491	Q492	F493	A494	S495	L496	E497	H498	G499	I500	V501	P502	V503	T504	S505	D506	C507	Q508	Y509	L510	F511	P512	A513	K514	V515	V516	S517	R518	L519	V520	K521	W522	V523	T524	V525	A526	H527	E528	D529	V530	M531	E532	L533	H534	F535	T536	L479	K480



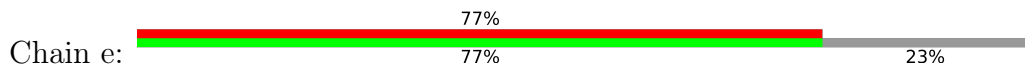
• Molecule 4: THO complex subunit 5 homolog



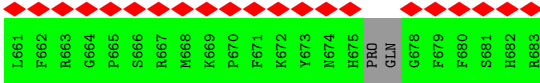
M601	M602	G603	E604	V605	M606	V607	C608	Y609	K610	E611	L612	C613	G614	P615	M616	P617	S618	H619	Q620	L621	L622	T623	N624	Q625	L626	Q627	R628	L629	C630	V631	L632	L633	D634	V635	Y636	L637	E638	T639	E640	S641	H642	ASP	ASP	SER	VAL	GLU	GLY	PRO	LYS	PHE	PRO	GLN	GLU	LYS	MET	CYS	L659	R660
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------

L661	F662	R663	G664	P665	S666	R667	M668	K669	P670	F671	K672	Y673	M674	H675	P676	Q677	G678	F679	F680	S681	H682	R683
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

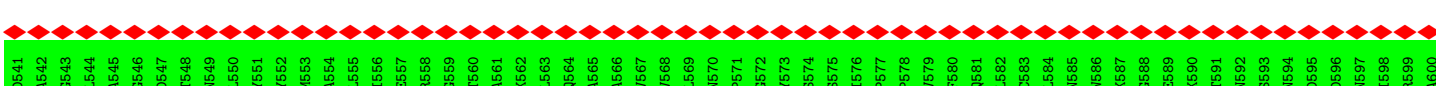
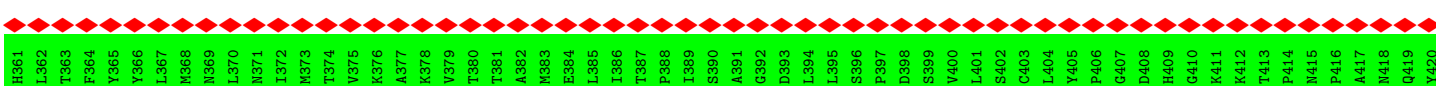
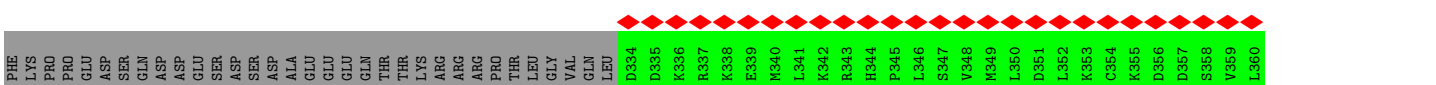
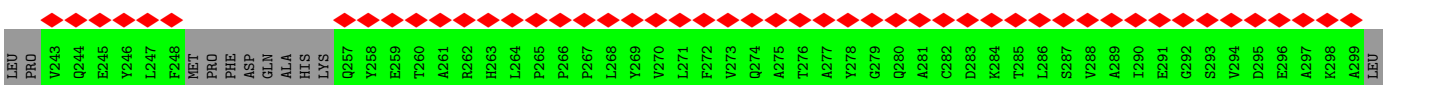
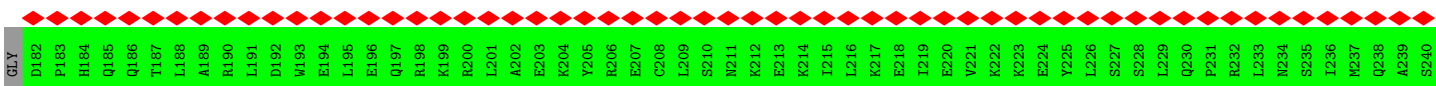
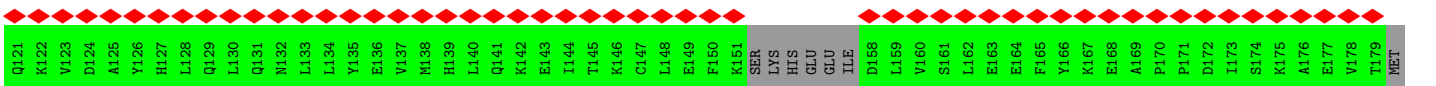
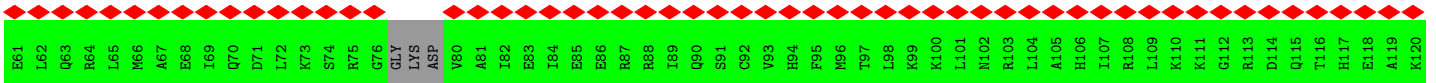
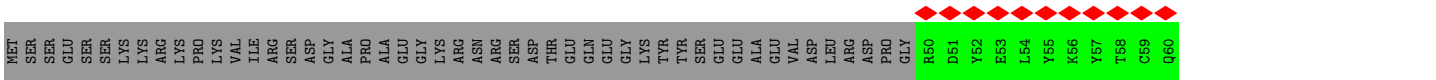
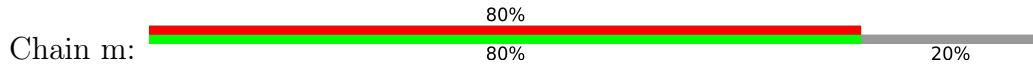
• Molecule 4: THO complex subunit 5 homolog



MET	SER	SER	GLU	SER	SER	LYS	LYS	ARG	LYS	PRO	LYS	VAL	ILE	ARG	ASP	GLY	ALA	PRO	GLY	ALA	GLY	GLY	GLY	GLY	GLY	GLY	TYR	S38	E39	E40	A41	E42	V43	D44	L45	R46	D47	PRO	GLY	ARG	ASP	Y52	E53	Y55	L54	Y56	K56	Y57	T58	C59	Q60										
E61	L62	Q63	R64	L65	M66	A67	E68	I69	Q70	D71	L72	K73	S74	R75	GLY	GLY	LYS	ASP	V80	A81	I82	E83	I84	E85	E86	R87	R88	I89	Q90	S91	C92	V93	H94	F95	M96	T97	L98	K99	K100	L101	M102	R103	L104	A105	H106	Y107	I108	L109	K110	K111	G112	R113	D114	Q115	T116	H117	E118	A119	K120		
Q121	K122	V123	D124	A125	Y126	H127	L128	Q129	L130	Q131	M132	L133	L134	Y135	E136	M137	H138	H139	L140	Q141	K142	E143	I144	THR	LYS	CYS	LEU	LEU	PHE	LYS	HIS	GLU	ILE	ASP	VAL	LEU	SER	LEU	GLU	GLU	PHE	TYR	LYS	ALA	PRO	PRO	ASP	ILE	SER	LYS	ALA	VAL	THR	MET							
GLY	ASP	PRO	HIS	GLN	THR	LEU	A189	M190	L191	D192	M193	E194	L195	E196	Q197	M198	K199	R200	L201	A202	E203	K204	Y205	R206	E207	C208	L209	S210	N211	K212	K214	L215	L216	L217	E218	I219	E220	V221	K222	K223	E224	V225	L226	S227	S228	L229	Q230	P231	K232	L233	N234	S235	T236	M237	Q238	A239	S240				
L241	P242	V243	Q244	E245	Y246	L247	PHE	MET	PRO	PHE	D252	Q253	A254	H255	K256	Q257	Y258	E259	T260	A261	R262	H263	L264	P265	P266	L268	Y269	V270	L271	F272	V273	Q274	A275	T276	A277	Y278	G279	Q280	A281	C282	D283	K284	T285	L286	S287	V288	A289	L290	E291	G292	S293	V294	D295	E296	A297	K298	A299	LEU			
PHE	LYS	PRO	GLU	ASP	SER	GLN	ASP	ASP	GLU	SER	SER	ASP	ALA	GLU	GLU	GLN	THR	LYS	ARG	ARG	PRO	THR	LEU	GLY	VAL	G1333	D334	D335	K336	R337	K338	E339	M340	L341	K342	R343	H344	P345	D346	S347	V348	R349	L350	D351	L352	C354	P406	K353	C354	D408	H409	G410	K411	K412	T413	P414	M415	P416	A417	Q419	Y420
H361	L362	T363	F364	Y365	Y366	L367	M368	N369	L370	M371	I372	M373	T374	V375	K376	A377	K378	V379	T380	T381	A382	M383	E384	L385	L386	T387	P388	I389	S390	A391	G392	D393	L394	L395	S396	P397	D398	S399	V400	L401	S402	C403	L404	Y405	P406	G407	H408	H409	G410	K411	K412	T413	P414	M415	P416	A417	Q419	Y420			
Q421	F422	D423	K424	V425	GLY	ILE	LEU	THR	L430	S431	D432	Y433	V434	L435	E436	L437	G438	H439	P440	Y441	L442	W443	V444	Q445	K446	L447	G448	G449	L450	H451	F452	P453	LYS	GLU	GLN	PRO	GLN	GLN	THR	VAL	I462	A463	D464	H465	S466	L467	S468	A469	H471	M472	S473	E474	T474	T475	M476	K477	L478	L479	K480		
T481	R482	V483	Q484	S485	R486	L487	A488	L489	H490	K491	Q492	F493	A494	S495	L496	E497	H498	G499	I500	V501	P502	V503	T504	S505	D506	C507	Q508	Y509	L510	F511	P512	A513	K514	V515	V516	S517	R518	L519	V520	K521	W522	W523	T524	V525	V526	H527	E528	D529	M531	F530	T531	T474	T475	M476	K477	L478	L479	K480			
D541	A542	G543	L544	G545	D547	T548	N549	L550	M551	Y552	M553	A554	L555	I556	E557	R558	G559	T560	A561	K562	L563	Q564	A565	A566	V567	V568	L569	M570	P571	G572	Y573	S574	S575	I576	P577	P578	V579	F580	Q581	L582	C583	L584	M585	W586	K587	G588	E589	K590	T591	M592	S593	M594	D595	D596	M597	I598	R599	A600			
M601	E602	G603	E604	V605	M606	V607	C608	Y609	K610	E611	L612	C613	G614	P615	M616	P617	S618	H619	Q620	L621	L622	T623	N624	Q625	L626	Q627	R628	L629	C630	V631	L632	L633	D634	V635	Y636	L637	E638	T639	E640	S641	H642	ASP	ASP	SER	VAL	GLU	GLY	PRO	LYS	PHE	PRO	GLN	GLU	LYS	MET	CYS	L659	R660			

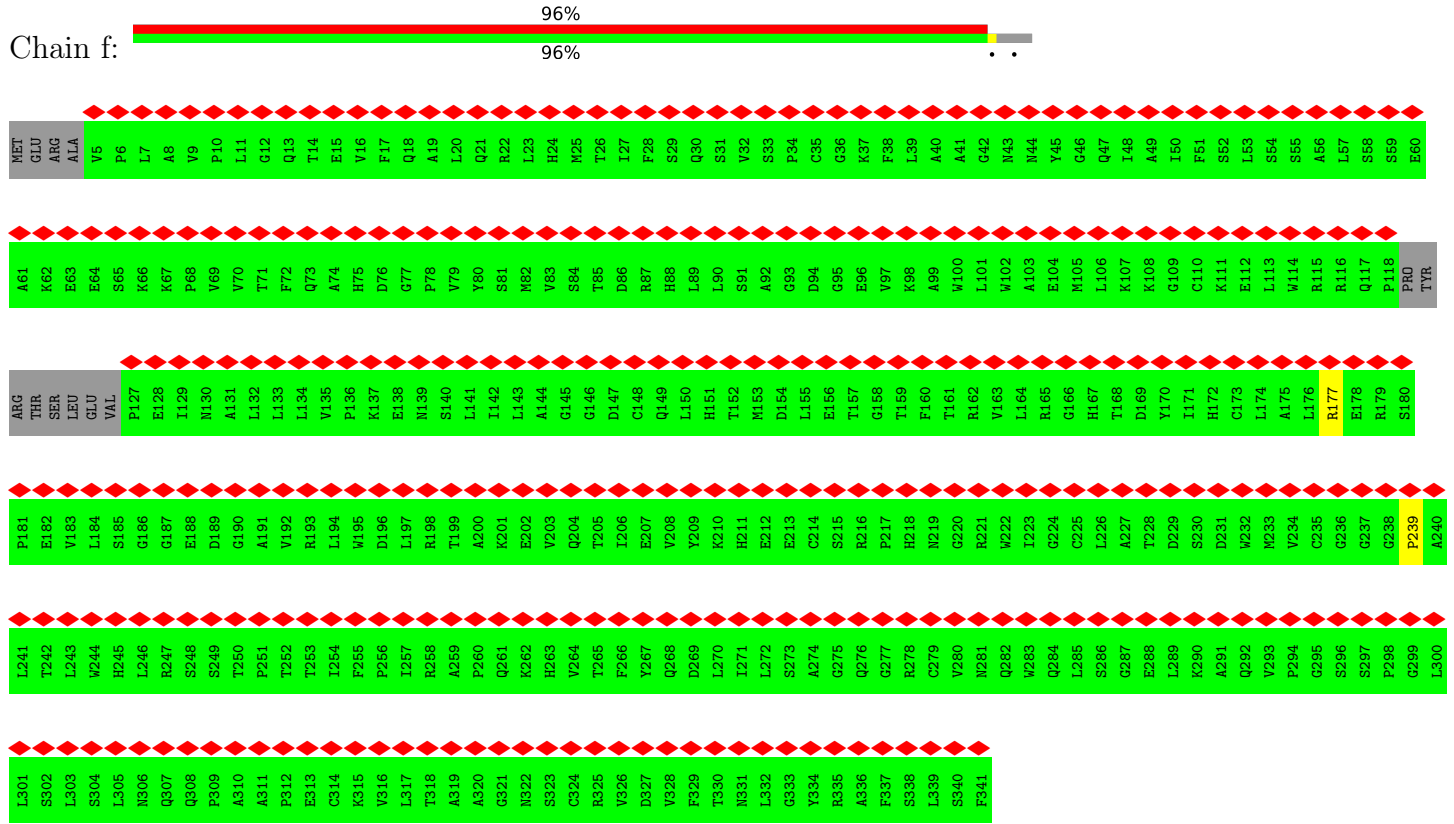


● Molecule 4: THO complex subunit 5 homolog

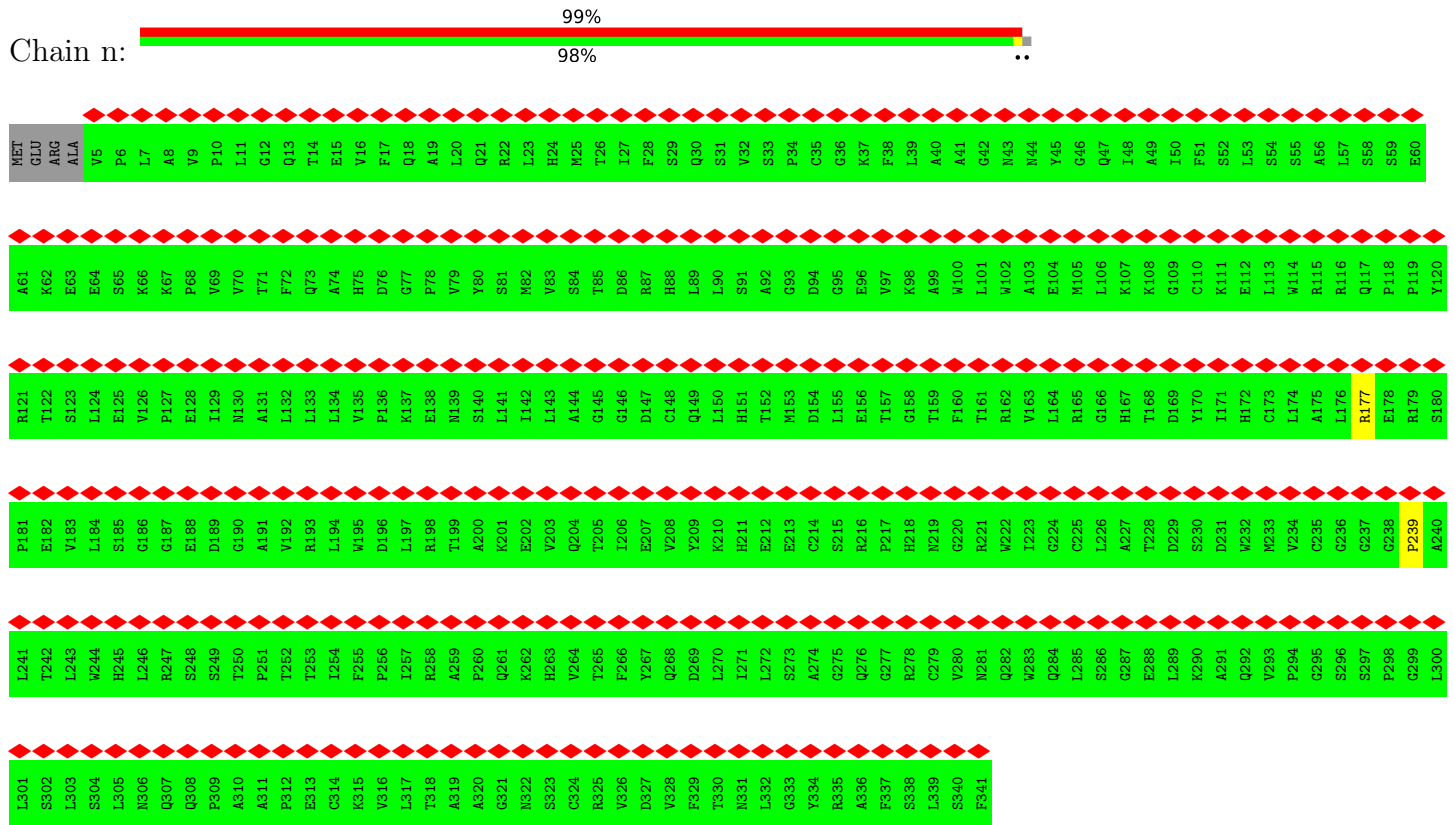




• Molecule 5: THO complex subunit 6 homolog

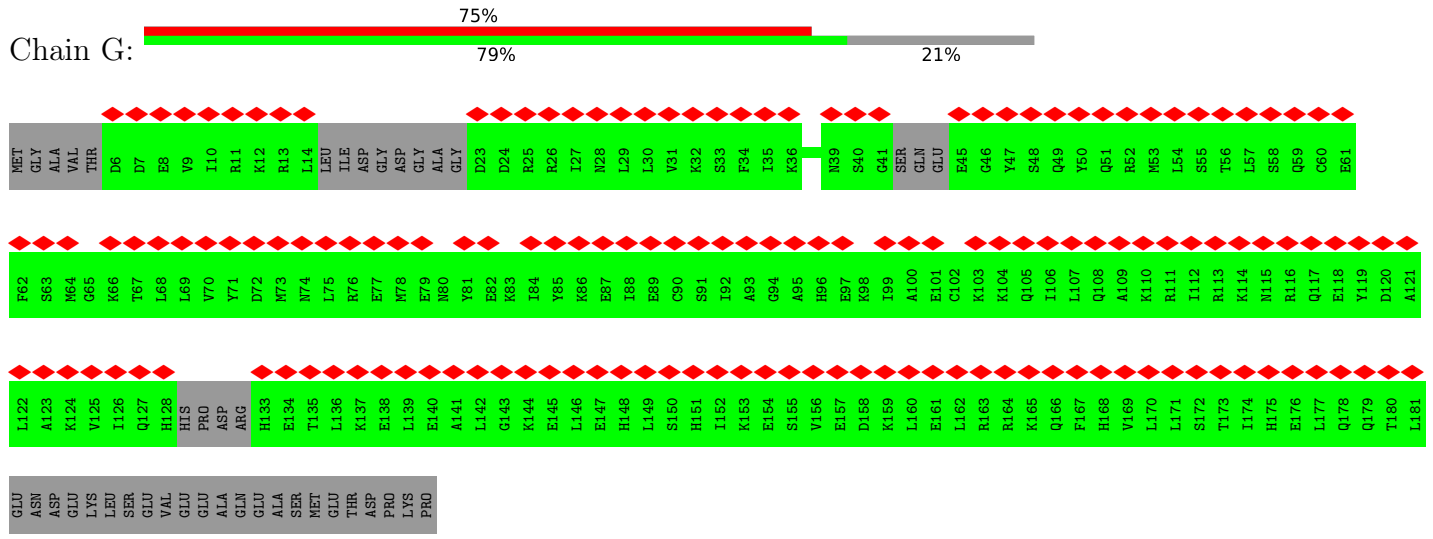


• Molecule 5: THO complex subunit 6 homolog

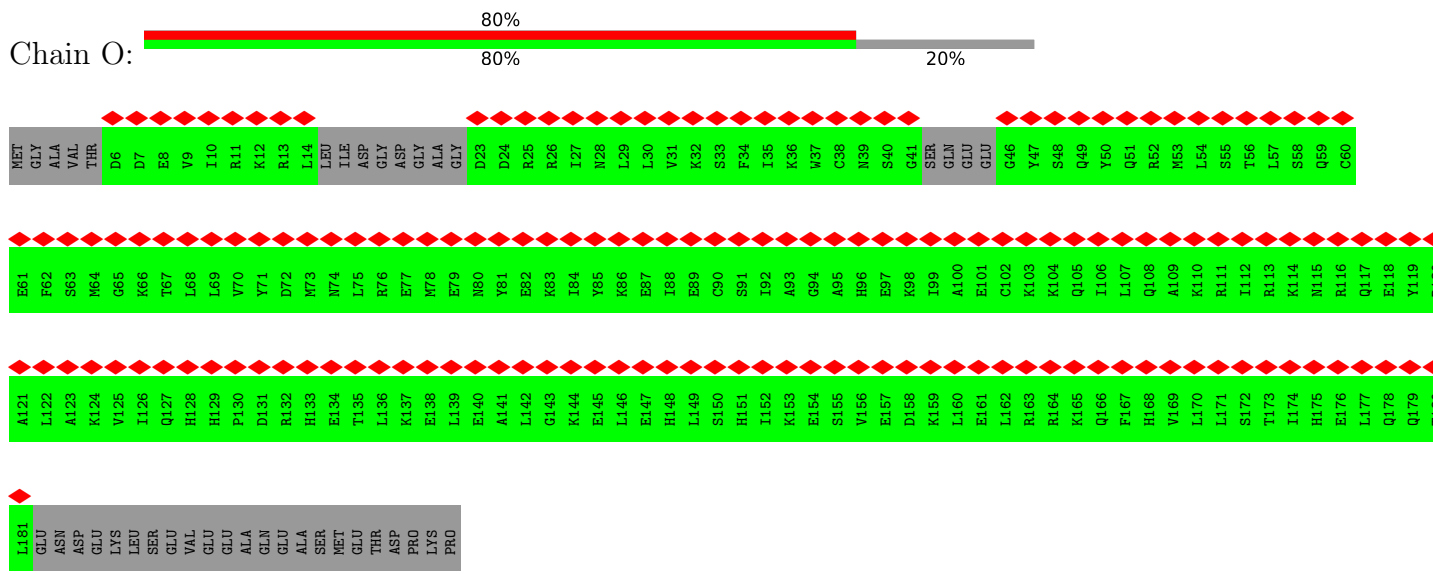




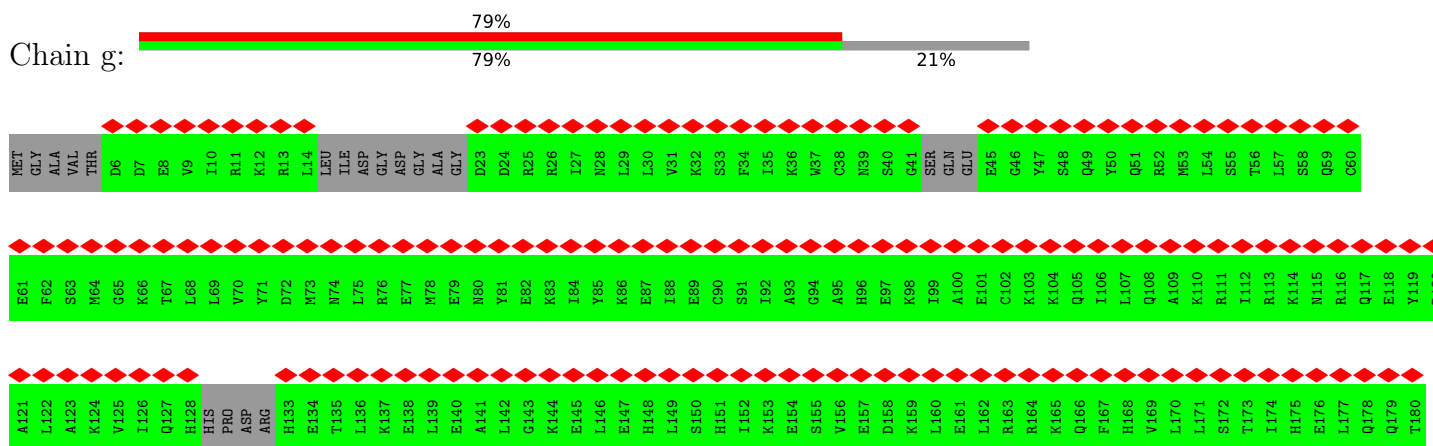
- Molecule 6: THO complex subunit 7 homolog



- Molecule 6: THO complex subunit 7 homolog



- Molecule 6: THO complex subunit 7 homolog





MET	HIS	ARG	LYS	T301	N361	S421
ALA	THR	ASN	PHE	A302	Y362	S422
GLU	ARG	LYS	MET	L303	D363	Y423
ASN	GLU	SER	GLN	A304	M364	I424
LEU	LEU	LEU	ASP	Q305	P365	E425
VAL	ALA	ASN	PRO	L306	E366	GLN
ASP	PHE	LEU	MET	L307	D367	THR
ASN	GLN	LYS	GLU	V308	S368	ARG
LEU	LEU	ILE	ILE	E309	D369	
LEU	SER	LYS	PHE	Q310	T370	
LEU	VAL	HIS	VAL	N311	Y371	
LEU	GLU	HIS	ASP	F312	L372	
LEU	GLU	PHE	ASP	P313	H373	
LEU	GLN	GLU	GLU	A314	R374	
LEU	VAL	CYS	VAL	I315	V375	
LEU	ALA	ASP	GLU	A316	A376	
LEU	THR	LYS	THR	I317	R377	
LEU	ALA	MET	ALA	H318	A378	
LEU	ALA	LEU	GLY	R319	G379	
LEU	GLY	GLN	GLY	G320	R380	
LEU	ASP	LYS	ASP	M321	F381	
LEU	GLY	GLN	GLY	P322	G382	
LEU	VAL	ARG	ALA	Q323	T383	
LEU	PRO	ASP	ALA	E324	K384	
LEU	LYS	VAL	VAL	E325	G385	
LEU	ALA	GLY	VAL	R326	L386	
LEU	SER	LYS	GLY	L327	A387	
LEU	LEU	ILE	VAL	S328	I388	
LEU	GLY	ASP	THR	R329	T389	
LEU	VAL	MET	VAL	F390	F390	
LEU	THR	THR	PHE	V391	V391	
LEU	GLU	PRO	VAL	Q332	S392	
LEU	VAL	HIS	VAL	F333	D393	
LEU	LEU	LYS	ALA	K334	E394	
LEU	ALA	LYS	THR	D335	N395	
LEU	ILE	GLN	LEU	F336	D396	
LEU	THR	VAL	GLY	Q337	A397	
LEU	THR	ARG	ARG	R338	K398	
LEU	THR	ALA	ASP	R339	I399	
LEU	THR	THR	PHE	L400	L400	
LEU	GLY	LEU	THR	N401	N401	
LEU	THR	SER	GLY	L341	L341	
LEU	GLN	LYS	LYS	V342	D402	
LEU	PRO	PRO	VAL	A343	V403	
LEU	VAL	GLU	VAL	T344	Q404	
LEU	VAL	ARG	VAL	N345	D405	
LEU	VAL	ILE	LEU	L346	R406	
LEU	VAL	ARG	LEU	F407	F407	
LEU	VAL	PRO	ALA	E408	E408	
LEU	ARG	VAL	VAL	V409	V409	
LEU	ALA	VAL	VAL	N410	N410	
LEU	VAL	VAL	VAL	I411	I411	
LEU	VAL	ARG	VAL	S412	S412	
LEU	VAL	ILE	VAL	E413	E413	
LEU	VAL	ARG	LEU	L414	L414	
LEU	VAL	PRO	LEU	P415	P415	
LEU	VAL	VAL	ALA	D416	D416	
LEU	VAL	VAL	CYS	E417	E417	
LEU	VAL	ARG	LEU	I418	I418	
LEU	VAL	ALA	LEU	D419	D419	
LEU	ALA	ARG	ARG	F360	F360	
LEU	ILE	ARG	ARG			

● Molecule 7: Spliceosome RNA helicase DDX39B



MET	HIS	ARG	LYS	I301	N361	S421
ALA	THR	ASN	PHE	A302	Y362	S422
GLU	ARG	LYS	MET	L303	D363	Y423
ASN	GLU	SER	GLN	A304	M364	I424
LEU	LEU	LEU	ASP	Q305	P365	E425
VAL	ALA	ASN	PRO	L306	E366	GLN
ASP	PHE	LEU	MET	L307	D367	THR
ASN	GLN	LYS	GLU	V308	S368	ARG
LEU	LEU	ILE	ILE	E309	D369	
LEU	SER	LYS	PHE	Q310	T370	
LEU	VAL	HIS	VAL	N311	Y371	
LEU	GLU	HIS	ASP	F312	L372	
LEU	GLU	PHE	ASP	P313	H373	
LEU	GLN	GLU	GLU	A314	R374	
LEU	VAL	CYS	VAL	I315	V375	
LEU	ALA	ASP	GLU	A316	A376	
LEU	THR	LYS	THR	I317	R377	
LEU	ALA	MET	ALA	H318	A378	
LEU	ALA	LEU	GLY	R319	G379	
LEU	GLY	GLN	GLY	G320	R380	
LEU	ASP	LYS	ASP	M321	F381	
LEU	VAL	ARG	ALA	P322	G382	
LEU	PRO	ASP	ALA	Q323	T383	
LEU	LYS	VAL	VAL	E324	K384	
LEU	ALA	GLY	VAL	E325	G385	
LEU	SER	LYS	GLY	R326	L386	
LEU	LEU	ILE	VAL	L327	A387	
LEU	THR	THR	THR	S328	I388	
LEU	GLU	ASP	ASP	R329	T389	
LEU	GLU	GLU	THR	F390	F390	
LEU	VAL	PRO	PHE	V391	V391	
LEU	VAL	HIS	VAL	Q331	S392	
LEU	VAL	LYS	VAL	Q332	S392	
LEU	LEU	LYS	ALA	F333	D393	
LEU	ALA	LYS	THR	K334	E394	
LEU	THR	ASN	LEU	D335	N395	
LEU	CYS	CYS	GLN	F336	D396	
LEU	PRO	HIS	GLN	Q337	A397	
LEU	ILE	PRO	PHE	R338	K398	
LEU	GLU	ILE	LEU	R339	I399	
LEU	ARG	ALA	VAL	L400	L400	
LEU	VAL	THR	VAL	N401	N401	
LEU	VAL	THR	THR	L341	L341	
LEU	GLY	LEU	GLY	V342	D402	
LEU	PRO	PRO	VAL	A343	V403	
LEU	VAL	ARG	VAL	T344	Q404	
LEU	VAL	LEU	VAL	N345	D405	
LEU	VAL	LEU	VAL	L346	R406	
LEU	VAL	ALA	VAL	F407	F407	
LEU	VAL	VAL	VAL	E408	E408	
LEU	VAL	VAL	VAL	V409	V409	
LEU	VAL	VAL	VAL	N410	N410	
LEU	VAL	VAL	VAL	I411	I411	
LEU	VAL	ARG	VAL	S412	S412	
LEU	VAL	ILE	VAL	E413	E413	
LEU	VAL	ARG	LEU	L414	L414	
LEU	VAL	PRO	LEU	P415	P415	
LEU	VAL	VAL	ALA	D416	D416	
LEU	VAL	VAL	CYS	E417	E417	
LEU	VAL	ARG	LEU	I418	I418	
LEU	VAL	ALA	LEU	D419	D419	
LEU	ALA	ARG	ARG	F360	F360	
LEU	ILE	ARG	ARG			



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	246457	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	50	Depositor
Minimum defocus (nm)	400	Depositor
Maximum defocus (nm)	3700	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	6.605	Depositor
Minimum map value	-4.523	Depositor
Average map value	0.003	Depositor
Map value standard deviation	0.048	Depositor
Recommended contour level	0.9	Depositor
Map size (Å)	589.60004, 589.60004, 589.60004	wwPDB
Map dimensions	440, 440, 440	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.34, 1.34, 1.34	Depositor

## 5 Model quality i

### 5.1 Standard geometry i

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	A	0.26	0/2942	0.38	0/3969
1	I	0.28	0/2989	0.46	0/4034
1	a	0.26	0/2942	0.38	0/3969
1	i	0.28	0/2990	0.45	0/4035
2	B	0.28	0/7129	0.43	1/9633 (0.0%)
2	J	0.28	1/7071 (0.0%)	0.43	1/9556 (0.0%)
2	b	0.28	0/7129	0.43	1/9633 (0.0%)
2	j	0.28	1/7071 (0.0%)	0.43	1/9556 (0.0%)
3	C	0.32	0/2494	0.49	0/3384
3	K	0.32	0/2494	0.49	0/3384
3	c	0.32	0/2494	0.49	0/3384
3	k	0.31	0/2494	0.49	0/3384
4	E	0.34	1/3830 (0.0%)	0.49	0/5225
4	M	0.29	0/4215	0.49	0/5717
4	e	0.34	1/3737 (0.0%)	0.49	0/5098
4	m	0.29	0/4215	0.49	0/5717
5	F	0.32	0/2666	0.57	0/3623
5	N	0.32	0/2666	0.56	0/3623
5	f	0.33	0/2596	0.57	0/3524
5	n	0.32	0/2666	0.57	0/3623
6	G	0.24	0/796	0.30	0/1105
6	O	0.27	0/1138	0.39	0/1536
6	g	0.24	0/796	0.30	0/1105
6	o	0.27	0/1138	0.39	0/1536
7	H	0.24	0/1421	0.39	0/1915
7	P	0.24	0/1421	0.39	0/1915
7	h	0.24	0/1421	0.39	0/1915
7	p	0.24	0/1421	0.39	0/1915
All	All	0.29	4/86382 (0.0%)	0.46	4/117013 (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
2	B	0	5
2	J	0	4
2	b	0	5
2	j	0	4
3	C	0	1
3	K	0	1
3	c	0	1
3	k	0	1
4	M	0	1
4	m	0	1
All	All	0	24

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	e	86	GLU	C-N	-10.02	1.11	1.34
4	E	86	GLU	C-N	-10.00	1.11	1.34
2	j	234	SER	C-N	-5.83	1.20	1.34
2	J	234	SER	C-N	-5.79	1.20	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	b	593	LEU	CA-CB-CG	6.33	129.86	115.30
2	J	593	LEU	CA-CB-CG	6.31	129.81	115.30
2	j	593	LEU	CA-CB-CG	6.31	129.80	115.30
2	B	593	LEU	CA-CB-CG	6.30	129.79	115.30

There are no chirality outliers.

All (24) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1116	GLU	Peptide
2	B	163	ASN	Peptide
2	B	557	THR	Peptide
2	B	562	LYS	Peptide
2	B	788	LEU	Peptide
3	C	287	ASP	Peptide
2	J	1116	GLU	Peptide
2	J	557	THR	Peptide
2	J	562	LYS	Peptide
2	J	788	LEU	Peptide

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
3	K	287	ASP	Peptide
4	M	660	ARG	Peptide
2	b	1116	GLU	Peptide
2	b	163	ASN	Peptide
2	b	557	THR	Peptide
2	b	562	LYS	Peptide
2	b	788	LEU	Peptide
3	c	287	ASP	Peptide
2	j	1116	GLU	Peptide
2	j	557	THR	Peptide
2	j	562	LYS	Peptide
2	j	788	LEU	Peptide
3	k	287	ASP	Peptide
4	m	660	ARG	Peptide

## 5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	330/657 (50%)	309 (94%)	19 (6%)	2 (1%)	25	62
1	I	335/657 (51%)	309 (92%)	23 (7%)	3 (1%)	17	54
1	a	330/657 (50%)	309 (94%)	19 (6%)	2 (1%)	25	62
1	i	335/657 (51%)	309 (92%)	23 (7%)	3 (1%)	17	54
2	B	871/1593 (55%)	805 (92%)	62 (7%)	4 (0%)	29	66
2	J	864/1593 (54%)	802 (93%)	59 (7%)	3 (0%)	41	75
2	b	871/1593 (55%)	805 (92%)	62 (7%)	4 (0%)	29	66
2	j	864/1593 (54%)	802 (93%)	59 (7%)	3 (0%)	41	75

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	305/351 (87%)	267 (88%)	37 (12%)	1 (0%)	41	75
3	K	305/351 (87%)	268 (88%)	36 (12%)	1 (0%)	41	75
3	c	305/351 (87%)	267 (88%)	37 (12%)	1 (0%)	41	75
3	k	305/351 (87%)	267 (88%)	37 (12%)	1 (0%)	41	75
4	E	520/683 (76%)	494 (95%)	25 (5%)	1 (0%)	47	80
4	M	529/683 (78%)	504 (95%)	25 (5%)	0	100	100
4	e	507/683 (74%)	485 (96%)	22 (4%)	0	100	100
4	m	529/683 (78%)	503 (95%)	26 (5%)	0	100	100
5	F	335/341 (98%)	305 (91%)	29 (9%)	1 (0%)	41	75
5	N	335/341 (98%)	306 (91%)	28 (8%)	1 (0%)	41	75
5	f	325/341 (95%)	298 (92%)	26 (8%)	1 (0%)	41	75
5	n	335/341 (98%)	305 (91%)	29 (9%)	1 (0%)	41	75
6	G	153/204 (75%)	153 (100%)	0	0	100	100
6	O	158/204 (78%)	158 (100%)	0	0	100	100
6	g	153/204 (75%)	153 (100%)	0	0	100	100
6	o	158/204 (78%)	158 (100%)	0	0	100	100
7	H	168/428 (39%)	162 (96%)	6 (4%)	0	100	100
7	P	168/428 (39%)	162 (96%)	6 (4%)	0	100	100
7	h	168/428 (39%)	162 (96%)	6 (4%)	0	100	100
7	p	168/428 (39%)	162 (96%)	6 (4%)	0	100	100
All	All	10729/17028 (63%)	9989 (93%)	707 (7%)	33 (0%)	44	75

All (33) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	789	SER
2	J	789	SER
2	b	789	SER
2	j	789	SER
2	B	557	THR
2	B	1116	GLU
1	I	308	MET
2	J	557	THR
2	J	1116	GLU
2	b	557	THR

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
2	b	1116	GLU
1	i	308	MET
2	j	557	THR
2	j	1116	GLU
2	B	164	LEU
1	I	105	PRO
2	b	164	LEU
1	i	105	PRO
3	C	31	PRO
3	K	31	PRO
3	c	31	PRO
3	k	31	PRO
4	E	409	HIS
5	F	239	PRO
5	N	239	PRO
5	f	239	PRO
5	n	239	PRO
1	A	105	PRO
1	A	434	GLY
1	I	434	GLY
1	a	105	PRO
1	a	434	GLY
1	i	434	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
1	A	310/601 (52%)	287 (93%)	23 (7%)	13 43
1	I	312/601 (52%)	307 (98%)	5 (2%)	62 83
1	a	310/601 (52%)	287 (93%)	23 (7%)	13 43
1	i	313/601 (52%)	308 (98%)	5 (2%)	62 83
2	B	704/1442 (49%)	632 (90%)	72 (10%)	7 30
2	J	697/1442 (48%)	632 (91%)	65 (9%)	9 34

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	b	704/1442 (49%)	633 (90%)	71 (10%)	7	30
2	j	697/1442 (48%)	632 (91%)	65 (9%)	9	34
3	C	268/300 (89%)	233 (87%)	35 (13%)	4	20
3	K	268/300 (89%)	233 (87%)	35 (13%)	4	20
3	c	268/300 (89%)	233 (87%)	35 (13%)	4	20
3	k	268/300 (89%)	233 (87%)	35 (13%)	4	20
4	E	322/615 (52%)	321 (100%)	1 (0%)	92	98
4	M	404/615 (66%)	404 (100%)	0	100	100
4	e	312/615 (51%)	312 (100%)	0	100	100
4	m	404/615 (66%)	404 (100%)	0	100	100
5	F	284/287 (99%)	283 (100%)	1 (0%)	91	97
5	N	284/287 (99%)	283 (100%)	1 (0%)	91	97
5	f	276/287 (96%)	275 (100%)	1 (0%)	91	97
5	n	284/287 (99%)	283 (100%)	1 (0%)	91	97
6	O	85/184 (46%)	85 (100%)	0	100	100
6	o	85/184 (46%)	85 (100%)	0	100	100
7	H	153/381 (40%)	153 (100%)	0	100	100
7	P	153/381 (40%)	153 (100%)	0	100	100
7	h	153/381 (40%)	153 (100%)	0	100	100
7	p	153/381 (40%)	153 (100%)	0	100	100
All	All	8471/14872 (57%)	7997 (94%)	474 (6%)	25	53

All (474) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	LEU
1	A	73	LEU
1	A	92	THR
1	A	142	ASP
1	A	152	ASN
1	A	234	LEU
1	A	251	CYS
1	A	271	PHE
1	A	301	PHE
1	A	323	LEU

Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	325	TYR
1	A	348	GLU
1	A	376	LYS
1	A	379	GLU
1	A	380	HIS
1	A	386	GLU
1	A	389	ASN
1	A	392	LYS
1	A	432	LEU
1	A	436	GLU
1	A	503	ARG
1	A	520	GLU
1	A	522	LEU
2	B	145	ASN
2	B	160	GLN
2	B	161	LYS
2	B	164	LEU
2	B	165	LEU
2	B	166	ARG
2	B	190	ASP
2	B	223	GLU
2	B	248	LEU
2	B	268	ARG
2	B	274	LEU
2	B	281	LEU
2	B	284	LEU
2	B	289	LEU
2	B	292	ASP
2	B	343	ASN
2	B	348	LEU
2	B	349	LEU
2	B	388	HIS
2	B	400	VAL
2	B	409	VAL
2	B	412	LEU
2	B	425	GLU
2	B	426	ASP
2	B	427	LEU
2	B	435	PHE
2	B	448	LEU
2	B	461	LYS
2	B	473	LYS

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	B	501	ASN
2	B	509	TRP
2	B	518	GLN
2	B	527	TRP
2	B	555	ARG
2	B	556	LEU
2	B	557	THR
2	B	591	ASP
2	B	593	LEU
2	B	594	ILE
2	B	631	ASP
2	B	633	THR
2	B	634	THR
2	B	642	LEU
2	B	657	LEU
2	B	660	LEU
2	B	678	ILE
2	B	701	GLU
2	B	729	LEU
2	B	749	GLN
2	B	766	VAL
2	B	800	ILE
2	B	801	ASP
2	B	812	ASP
2	B	879	GLN
2	B	880	PHE
2	B	892	LEU
2	B	938	LEU
2	B	939	LEU
2	B	953	LEU
2	B	956	LEU
2	B	978	LEU
2	B	987	ILE
2	B	996	CYS
2	B	1005	GLN
2	B	1034	ASN
2	B	1042	PHE
2	B	1044	CYS
2	B	1047	LEU
2	B	1101	HIS
2	B	1104	LEU
2	B	1116	GLU

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	B	1129	LYS
3	C	48	THR
3	C	52	LEU
3	C	61	VAL
3	C	63	TRP
3	C	66	ASP
3	C	68	ARG
3	C	69	ARG
3	C	70	LEU
3	C	83	LEU
3	C	89	LEU
3	C	104	GLN
3	C	113	ASP
3	C	114	LEU
3	C	121	ASP
3	C	124	ILE
3	C	134	CYS
3	C	135	ILE
3	C	152	ASP
3	C	158	VAL
3	C	163	ASP
3	C	172	THR
3	C	185	VAL
3	C	192	ASN
3	C	199	LEU
3	C	206	ILE
3	C	213	GLU
3	C	226	ASN
3	C	260	ARG
3	C	281	LEU
3	C	301	LEU
3	C	303	GLU
3	C	304	VAL
3	C	312	THR
3	C	322	LEU
3	C	347	LEU
4	E	524	THR
5	F	177	ARG
1	I	432	LEU
1	I	436	GLU
1	I	503	ARG
1	I	520	GLU

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	I	522	LEU
2	J	145	ASN
2	J	204	ASN
2	J	248	LEU
2	J	268	ARG
2	J	274	LEU
2	J	281	LEU
2	J	284	LEU
2	J	289	LEU
2	J	292	ASP
2	J	343	ASN
2	J	348	LEU
2	J	349	LEU
2	J	388	HIS
2	J	400	VAL
2	J	409	VAL
2	J	412	LEU
2	J	425	GLU
2	J	426	ASP
2	J	427	LEU
2	J	435	PHE
2	J	448	LEU
2	J	461	LYS
2	J	473	LYS
2	J	501	ASN
2	J	509	TRP
2	J	518	GLN
2	J	527	TRP
2	J	555	ARG
2	J	556	LEU
2	J	557	THR
2	J	591	ASP
2	J	593	LEU
2	J	594	ILE
2	J	631	ASP
2	J	633	THR
2	J	634	THR
2	J	642	LEU
2	J	657	LEU
2	J	660	LEU
2	J	678	ILE
2	J	729	LEU

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	J	749	GLN
2	J	766	VAL
2	J	800	ILE
2	J	801	ASP
2	J	812	ASP
2	J	879	GLN
2	J	880	PHE
2	J	892	LEU
2	J	938	LEU
2	J	939	LEU
2	J	953	LEU
2	J	956	LEU
2	J	978	LEU
2	J	987	ILE
2	J	996	CYS
2	J	1005	GLN
2	J	1034	ASN
2	J	1042	PHE
2	J	1044	CYS
2	J	1047	LEU
2	J	1101	HIS
2	J	1104	LEU
2	J	1116	GLU
2	J	1129	LYS
3	K	48	THR
3	K	52	LEU
3	K	61	VAL
3	K	63	TRP
3	K	66	ASP
3	K	68	ARG
3	K	69	ARG
3	K	70	LEU
3	K	83	LEU
3	K	89	LEU
3	K	104	GLN
3	K	113	ASP
3	K	114	LEU
3	K	121	ASP
3	K	124	ILE
3	K	134	CYS
3	K	135	ILE
3	K	152	ASP

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
3	K	158	VAL
3	K	163	ASP
3	K	172	THR
3	K	185	VAL
3	K	192	ASN
3	K	199	LEU
3	K	206	ILE
3	K	213	GLU
3	K	226	ASN
3	K	260	ARG
3	K	281	LEU
3	K	301	LEU
3	K	303	GLU
3	K	304	VAL
3	K	312	THR
3	K	322	LEU
3	K	347	LEU
5	N	177	ARG
1	a	51	LEU
1	a	73	LEU
1	a	92	THR
1	a	142	ASP
1	a	152	ASN
1	a	234	LEU
1	a	251	CYS
1	a	271	PHE
1	a	301	PHE
1	a	323	LEU
1	a	325	TYR
1	a	348	GLU
1	a	376	LYS
1	a	379	GLU
1	a	380	HIS
1	a	386	GLU
1	a	389	ASN
1	a	392	LYS
1	a	432	LEU
1	a	436	GLU
1	a	503	ARG
1	a	520	GLU
1	a	522	LEU
2	b	145	ASN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	b	160	GLN
2	b	161	LYS
2	b	164	LEU
2	b	165	LEU
2	b	166	ARG
2	b	190	ASP
2	b	223	GLU
2	b	248	LEU
2	b	268	ARG
2	b	274	LEU
2	b	281	LEU
2	b	284	LEU
2	b	289	LEU
2	b	292	ASP
2	b	343	ASN
2	b	348	LEU
2	b	349	LEU
2	b	388	HIS
2	b	400	VAL
2	b	409	VAL
2	b	412	LEU
2	b	425	GLU
2	b	426	ASP
2	b	427	LEU
2	b	435	PHE
2	b	448	LEU
2	b	461	LYS
2	b	473	LYS
2	b	501	ASN
2	b	509	TRP
2	b	518	GLN
2	b	527	TRP
2	b	555	ARG
2	b	556	LEU
2	b	557	THR
2	b	591	ASP
2	b	593	LEU
2	b	594	ILE
2	b	631	ASP
2	b	633	THR
2	b	634	THR
2	b	642	LEU

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	b	657	LEU
2	b	660	LEU
2	b	678	ILE
2	b	729	LEU
2	b	749	GLN
2	b	766	VAL
2	b	800	ILE
2	b	801	ASP
2	b	812	ASP
2	b	879	GLN
2	b	880	PHE
2	b	892	LEU
2	b	938	LEU
2	b	939	LEU
2	b	953	LEU
2	b	956	LEU
2	b	978	LEU
2	b	987	ILE
2	b	996	CYS
2	b	1005	GLN
2	b	1034	ASN
2	b	1042	PHE
2	b	1044	CYS
2	b	1047	LEU
2	b	1101	HIS
2	b	1104	LEU
2	b	1116	GLU
2	b	1129	LYS
3	c	48	THR
3	c	52	LEU
3	c	61	VAL
3	c	63	TRP
3	c	66	ASP
3	c	68	ARG
3	c	69	ARG
3	c	70	LEU
3	c	83	LEU
3	c	89	LEU
3	c	104	GLN
3	c	113	ASP
3	c	114	LEU
3	c	121	ASP

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
3	c	124	ILE
3	c	134	CYS
3	c	135	ILE
3	c	152	ASP
3	c	158	VAL
3	c	163	ASP
3	c	172	THR
3	c	185	VAL
3	c	192	ASN
3	c	199	LEU
3	c	206	ILE
3	c	213	GLU
3	c	226	ASN
3	c	260	ARG
3	c	281	LEU
3	c	301	LEU
3	c	303	GLU
3	c	304	VAL
3	c	312	THR
3	c	322	LEU
3	c	347	LEU
5	f	177	ARG
1	i	432	LEU
1	i	436	GLU
1	i	503	ARG
1	i	520	GLU
1	i	522	LEU
2	j	145	ASN
2	j	204	ASN
2	j	248	LEU
2	j	268	ARG
2	j	274	LEU
2	j	281	LEU
2	j	284	LEU
2	j	289	LEU
2	j	292	ASP
2	j	343	ASN
2	j	348	LEU
2	j	349	LEU
2	j	388	HIS
2	j	400	VAL
2	j	409	VAL

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	j	412	LEU
2	j	425	GLU
2	j	426	ASP
2	j	427	LEU
2	j	435	PHE
2	j	448	LEU
2	j	461	LYS
2	j	473	LYS
2	j	501	ASN
2	j	509	TRP
2	j	518	GLN
2	j	527	TRP
2	j	555	ARG
2	j	556	LEU
2	j	557	THR
2	j	591	ASP
2	j	593	LEU
2	j	594	ILE
2	j	631	ASP
2	j	633	THR
2	j	634	THR
2	j	642	LEU
2	j	657	LEU
2	j	660	LEU
2	j	678	ILE
2	j	729	LEU
2	j	749	GLN
2	j	766	VAL
2	j	800	ILE
2	j	801	ASP
2	j	812	ASP
2	j	879	GLN
2	j	880	PHE
2	j	892	LEU
2	j	938	LEU
2	j	939	LEU
2	j	953	LEU
2	j	956	LEU
2	j	978	LEU
2	j	987	ILE
2	j	996	CYS
2	j	1005	GLN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	j	1034	ASN
2	j	1042	PHE
2	j	1044	CYS
2	j	1047	LEU
2	j	1101	HIS
2	j	1104	LEU
2	j	1116	GLU
2	j	1129	LYS
3	k	48	THR
3	k	52	LEU
3	k	61	VAL
3	k	63	TRP
3	k	66	ASP
3	k	68	ARG
3	k	69	ARG
3	k	70	LEU
3	k	83	LEU
3	k	89	LEU
3	k	104	GLN
3	k	113	ASP
3	k	114	LEU
3	k	121	ASP
3	k	124	ILE
3	k	134	CYS
3	k	135	ILE
3	k	152	ASP
3	k	158	VAL
3	k	163	ASP
3	k	172	THR
3	k	185	VAL
3	k	192	ASN
3	k	199	LEU
3	k	206	ILE
3	k	213	GLU
3	k	226	ASN
3	k	260	ARG
3	k	281	LEU
3	k	301	LEU
3	k	303	GLU
3	k	304	VAL
3	k	312	THR
3	k	322	LEU

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
3	k	347	LEU
5	n	177	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (189) such sidechains are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	160	GLN
1	A	242	GLN
1	A	320	HIS
1	A	324	GLN
1	A	512	ASN
1	A	524	ASN
2	B	160	GLN
2	B	204	ASN
2	B	256	GLN
2	B	343	ASN
2	B	360	HIS
2	B	362	GLN
2	B	367	GLN
2	B	388	HIS
2	B	441	HIS
2	B	489	GLN
2	B	526	GLN
2	B	535	HIS
2	B	567	GLN
2	B	573	HIS
2	B	586	GLN
2	B	592	ASN
2	B	608	ASN
2	B	640	GLN
2	B	727	GLN
2	B	749	GLN
2	B	779	GLN
2	B	787	ASN
2	B	805	ASN
2	B	945	GLN
2	B	1005	GLN
3	C	59	HIS
3	C	104	GLN
3	C	146	ASN
3	C	195	ASN
3	C	203	ASN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
3	C	207	ASN
3	C	305	GLN
3	C	316	HIS
4	E	274	GLN
4	E	369	ASN
4	E	465	HIS
4	E	492	GLN
4	E	675	HIS
5	F	130	ASN
5	F	149	GLN
5	F	172	HIS
5	F	261	GLN
7	H	331	GLN
7	H	404	GLN
1	I	141	ASN
1	I	242	GLN
1	I	512	ASN
1	I	524	ASN
2	J	204	ASN
2	J	224	HIS
2	J	256	GLN
2	J	343	ASN
2	J	360	HIS
2	J	362	GLN
2	J	367	GLN
2	J	388	HIS
2	J	441	HIS
2	J	489	GLN
2	J	526	GLN
2	J	535	HIS
2	J	567	GLN
2	J	573	HIS
2	J	586	GLN
2	J	592	ASN
2	J	608	ASN
2	J	640	GLN
2	J	727	GLN
2	J	749	GLN
2	J	779	GLN
2	J	787	ASN
2	J	805	ASN
2	J	945	GLN

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	J	1005	GLN
3	K	59	HIS
3	K	104	GLN
3	K	146	ASN
3	K	195	ASN
3	K	203	ASN
3	K	207	ASN
3	K	305	GLN
3	K	316	HIS
4	M	369	ASN
4	M	492	GLN
4	M	675	HIS
5	N	149	GLN
5	N	172	HIS
5	N	261	GLN
7	P	331	GLN
7	P	404	GLN
1	a	160	GLN
1	a	242	GLN
1	a	320	HIS
1	a	324	GLN
1	a	512	ASN
1	a	524	ASN
2	b	160	GLN
2	b	204	ASN
2	b	256	GLN
2	b	343	ASN
2	b	360	HIS
2	b	362	GLN
2	b	367	GLN
2	b	388	HIS
2	b	441	HIS
2	b	489	GLN
2	b	526	GLN
2	b	535	HIS
2	b	567	GLN
2	b	573	HIS
2	b	586	GLN
2	b	592	ASN
2	b	608	ASN
2	b	640	GLN
2	b	727	GLN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	b	749	GLN
2	b	779	GLN
2	b	787	ASN
2	b	805	ASN
2	b	945	GLN
2	b	1005	GLN
3	c	59	HIS
3	c	104	GLN
3	c	146	ASN
3	c	195	ASN
3	c	203	ASN
3	c	207	ASN
3	c	305	GLN
3	c	316	HIS
4	e	274	GLN
4	e	369	ASN
4	e	465	HIS
4	e	492	GLN
4	e	675	HIS
5	f	130	ASN
5	f	172	HIS
5	f	261	GLN
7	h	331	GLN
7	h	404	GLN
1	i	141	ASN
1	i	242	GLN
1	i	512	ASN
1	i	524	ASN
2	j	204	ASN
2	j	224	HIS
2	j	256	GLN
2	j	343	ASN
2	j	360	HIS
2	j	362	GLN
2	j	367	GLN
2	j	388	HIS
2	j	441	HIS
2	j	489	GLN
2	j	526	GLN
2	j	535	HIS
2	j	567	GLN
2	j	573	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	j	586	GLN
2	j	592	ASN
2	j	608	ASN
2	j	640	GLN
2	j	727	GLN
2	j	749	GLN
2	j	779	GLN
2	j	787	ASN
2	j	805	ASN
2	j	945	GLN
2	j	1005	GLN
3	k	59	HIS
3	k	104	GLN
3	k	146	ASN
3	k	195	ASN
3	k	203	ASN
3	k	207	ASN
3	k	305	GLN
3	k	316	HIS
4	m	369	ASN
4	m	492	GLN
4	m	675	HIS
5	n	149	GLN
5	n	172	HIS
5	n	261	GLN
7	p	331	GLN
7	p	404	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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 [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	E	1
4	e	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	86:GLU	C	87:ARG	N	1.11
1	e	86:GLU	C	87:ARG	N	1.11

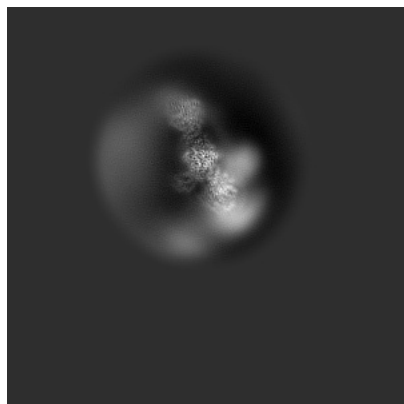
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-14808. These allow visual inspection of the internal detail of the map and identification of artifacts.

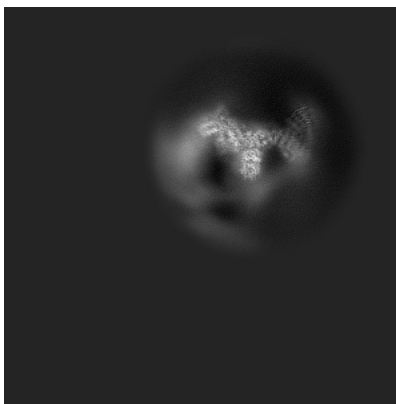
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections [i](#)

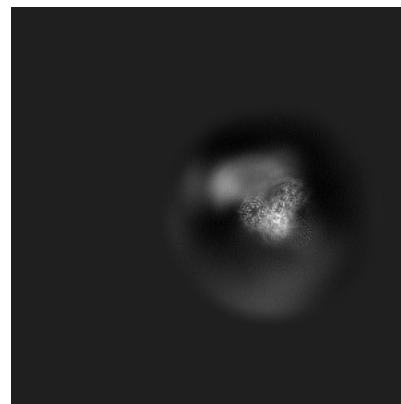
#### 6.1.1 Primary map



X

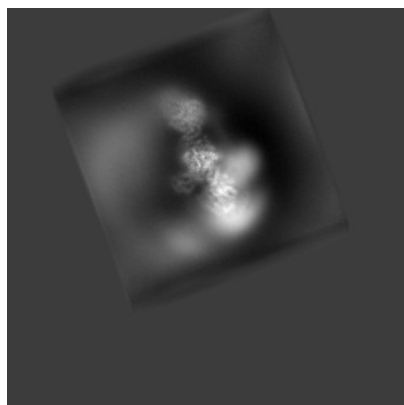


Y

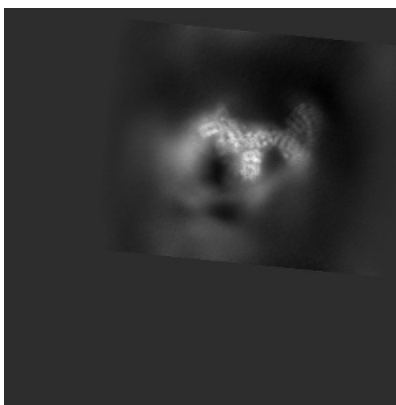


Z

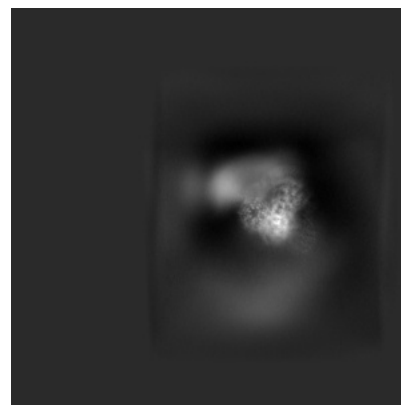
#### 6.1.2 Raw map



X



Y

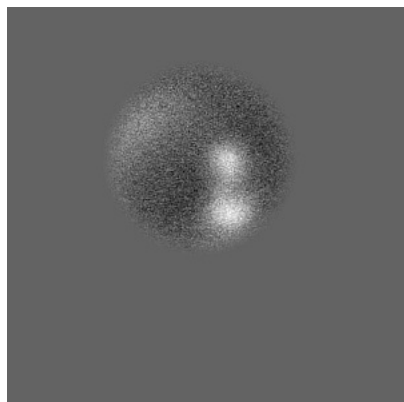


Z

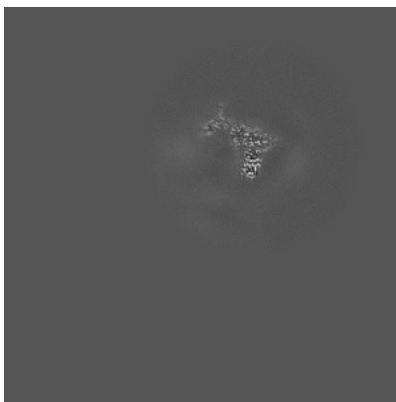
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

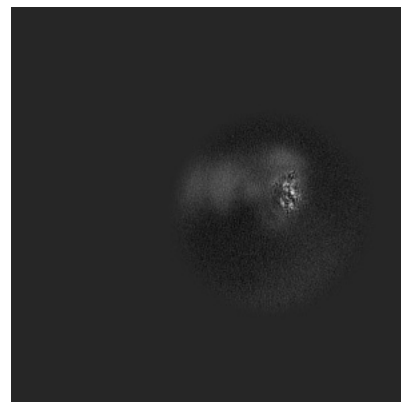
### 6.2.1 Primary map



X Index: 220

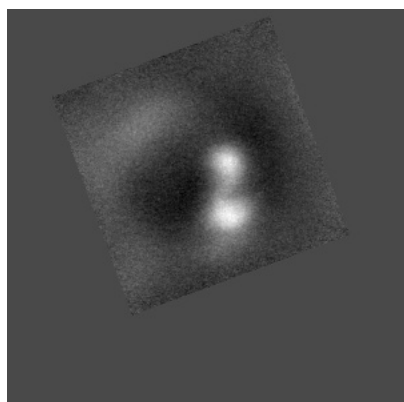


Y Index: 220

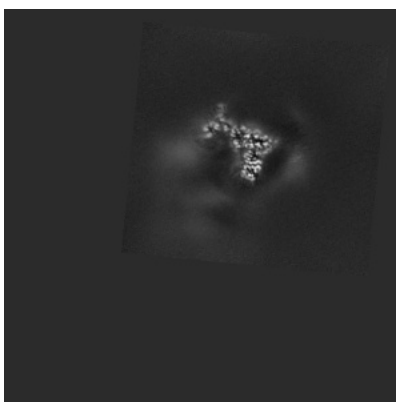


Z Index: 220

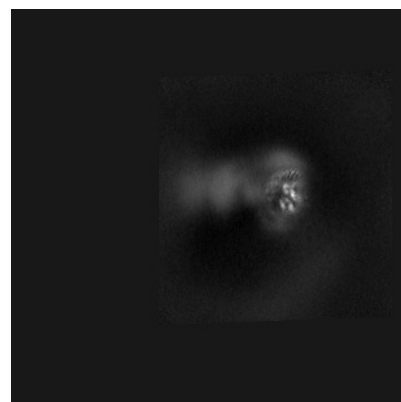
### 6.2.2 Raw map



X Index: 220



Y Index: 220

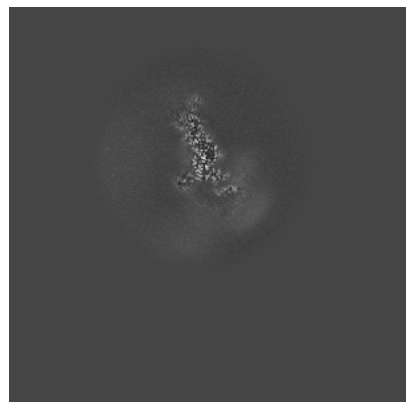


Z Index: 220

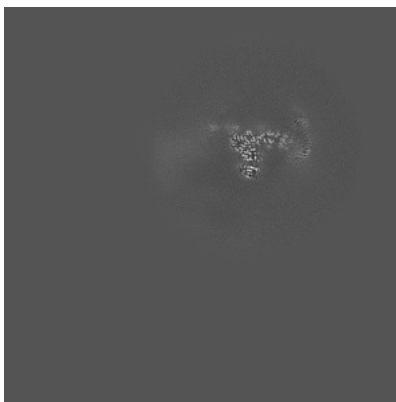
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

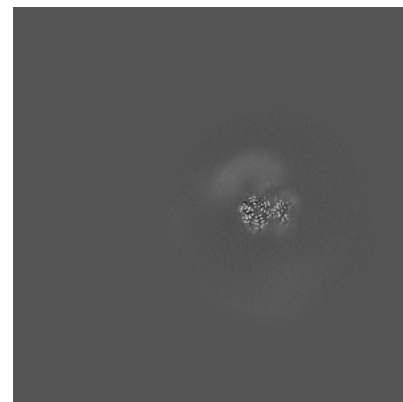
### 6.3.1 Primary map



X Index: 295

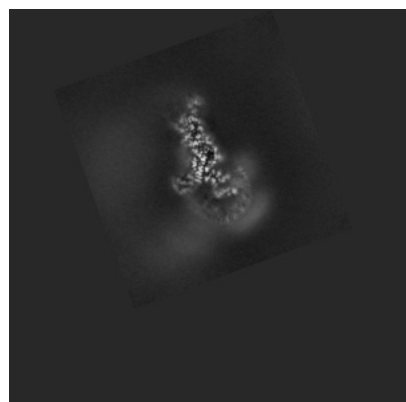


Y Index: 211

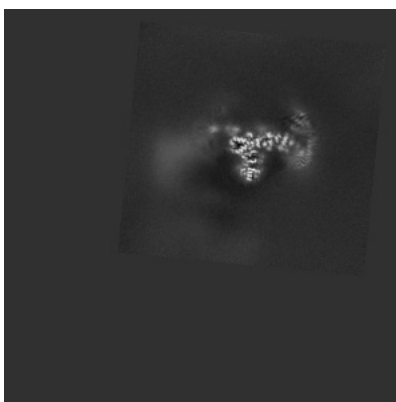


Z Index: 272

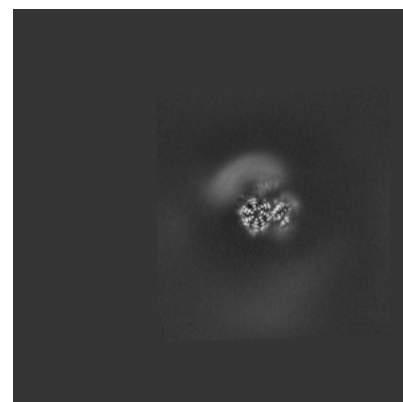
### 6.3.2 Raw map



X Index: 295



Y Index: 210

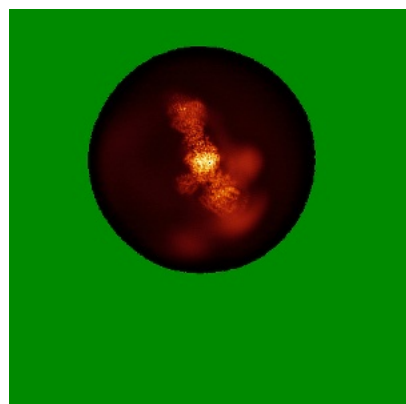


Z Index: 272

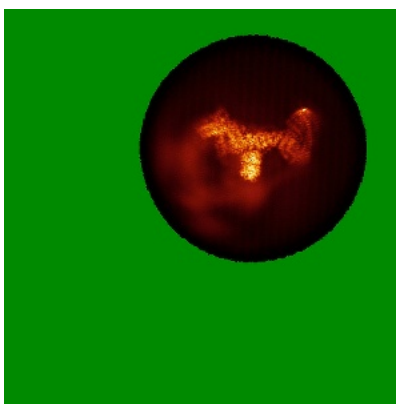
The images above show the largest variance slices of the map in three orthogonal directions.

## 6.4 Orthogonal standard-deviation projections (False-color) [i](#)

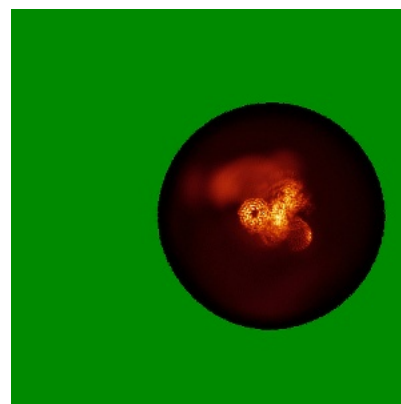
### 6.4.1 Primary map



X

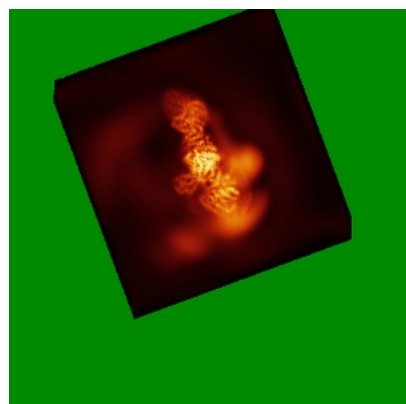


Y

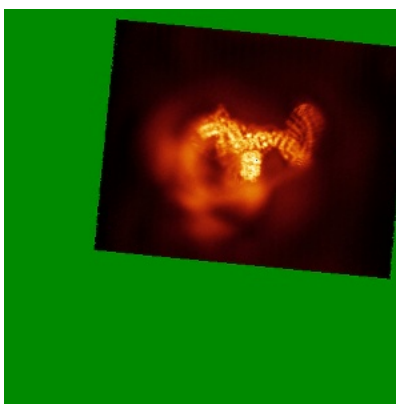


Z

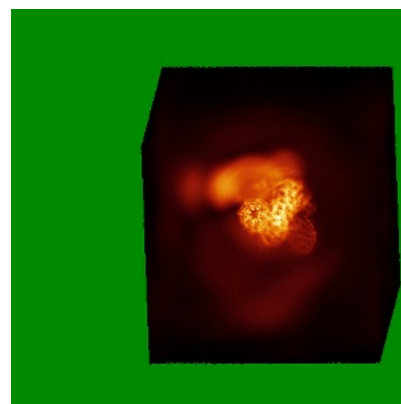
### 6.4.2 Raw map



X



Y



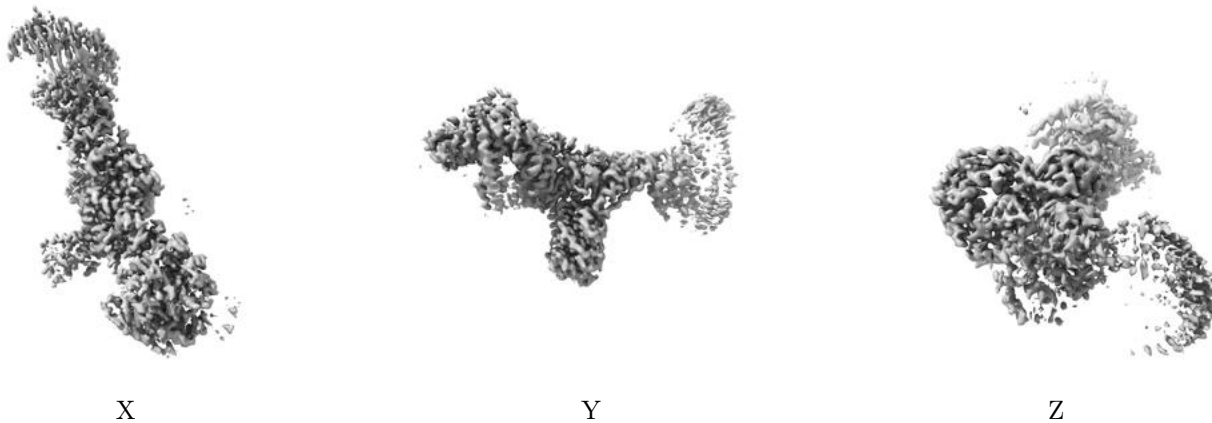
Z

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



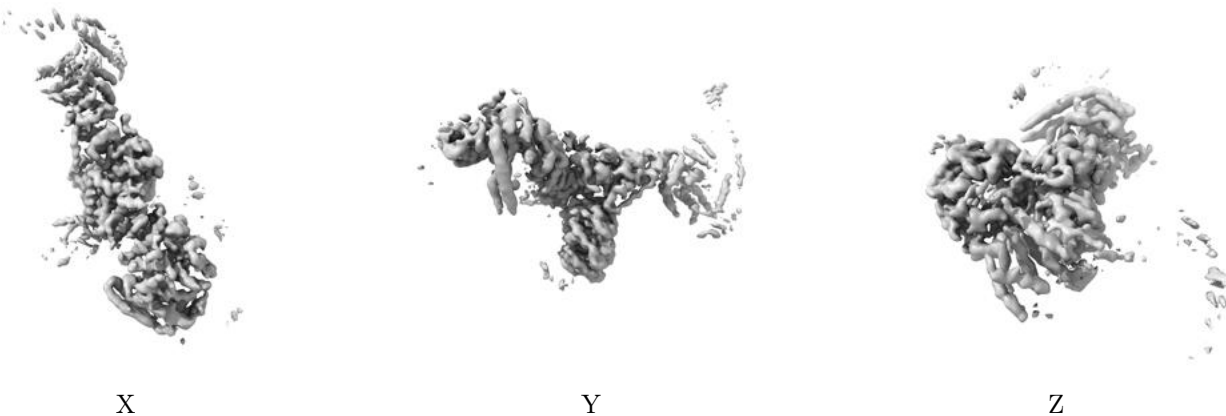
## 6.5 Orthogonal surface views [i](#)

### 6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.9. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

### 6.5.2 Raw map



These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

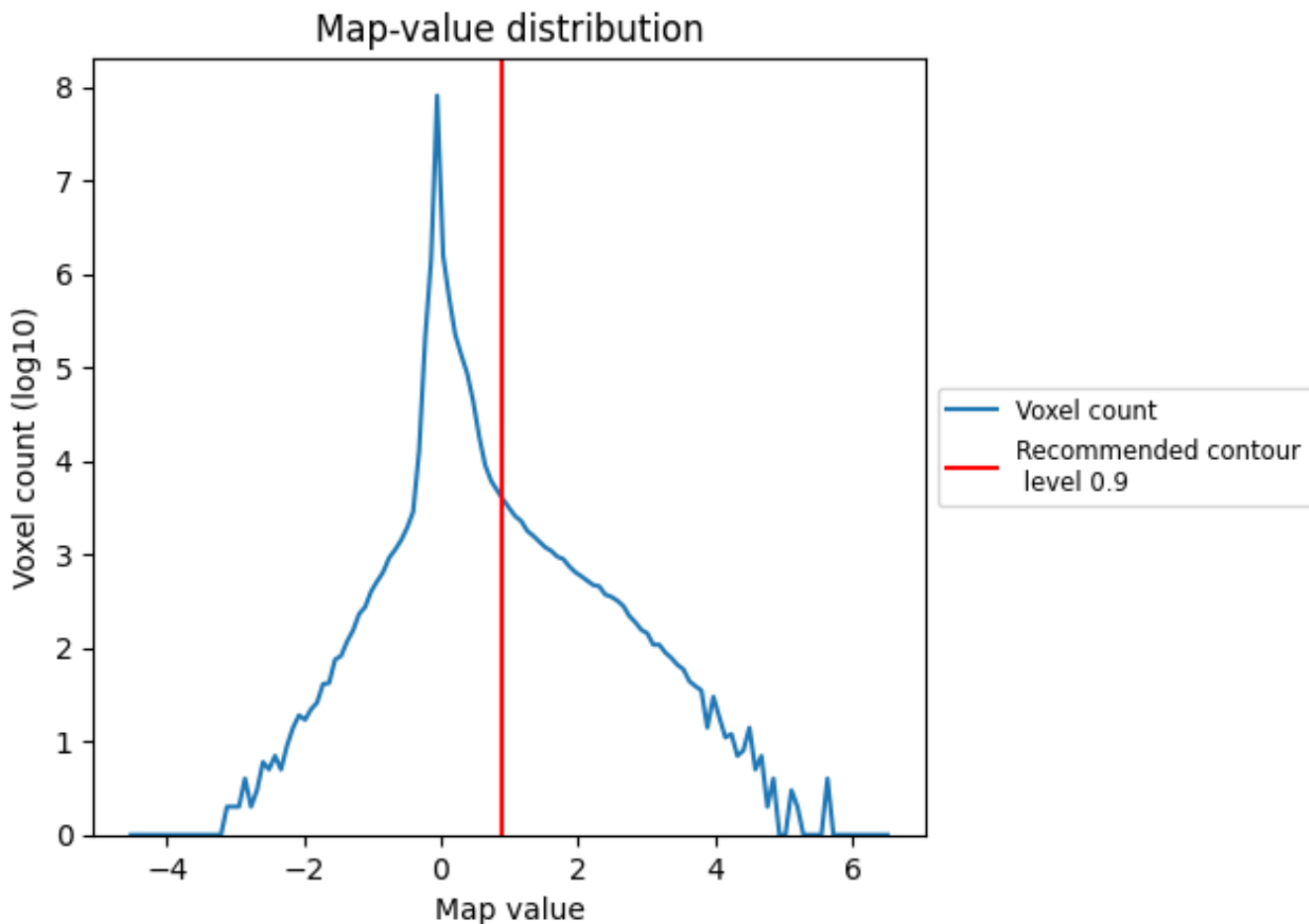
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

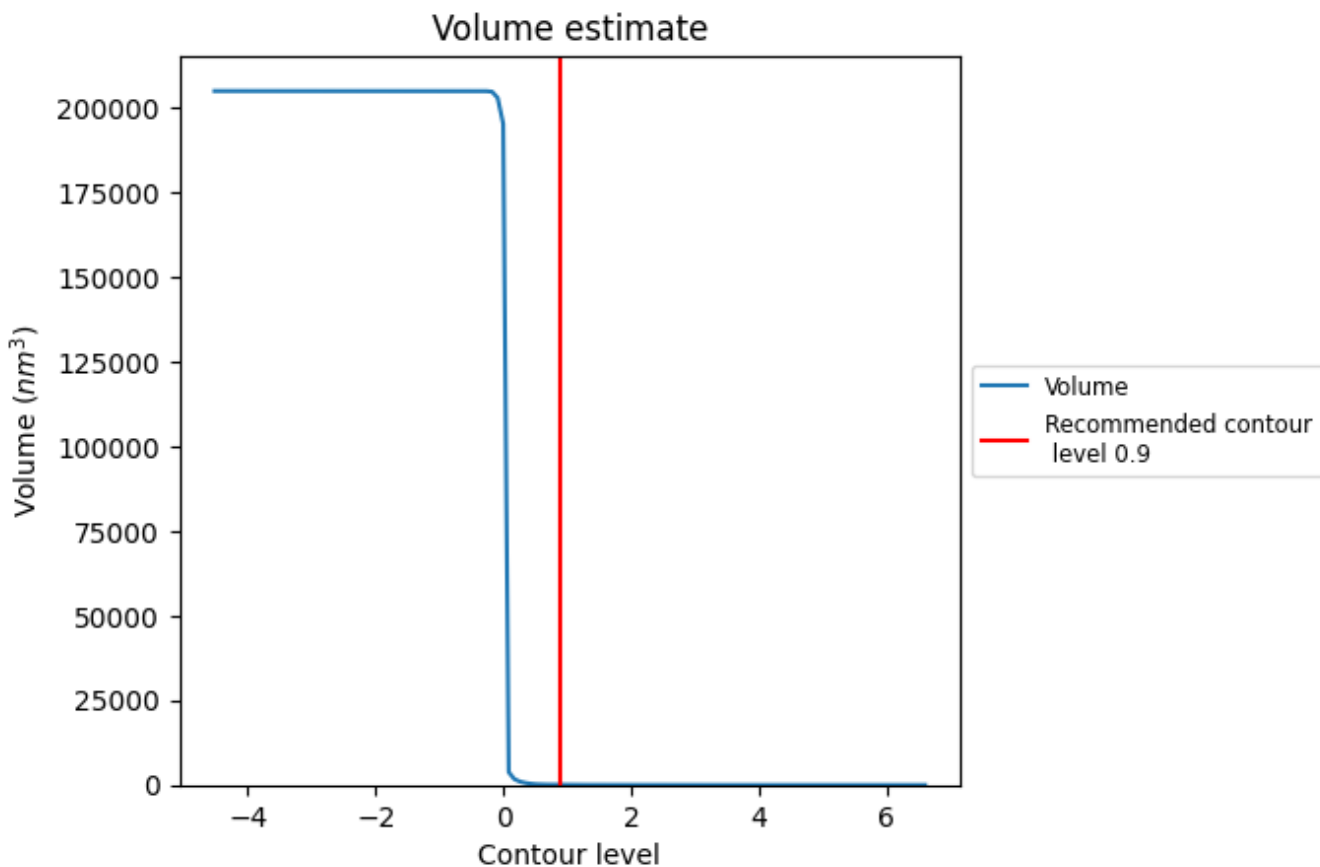
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



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

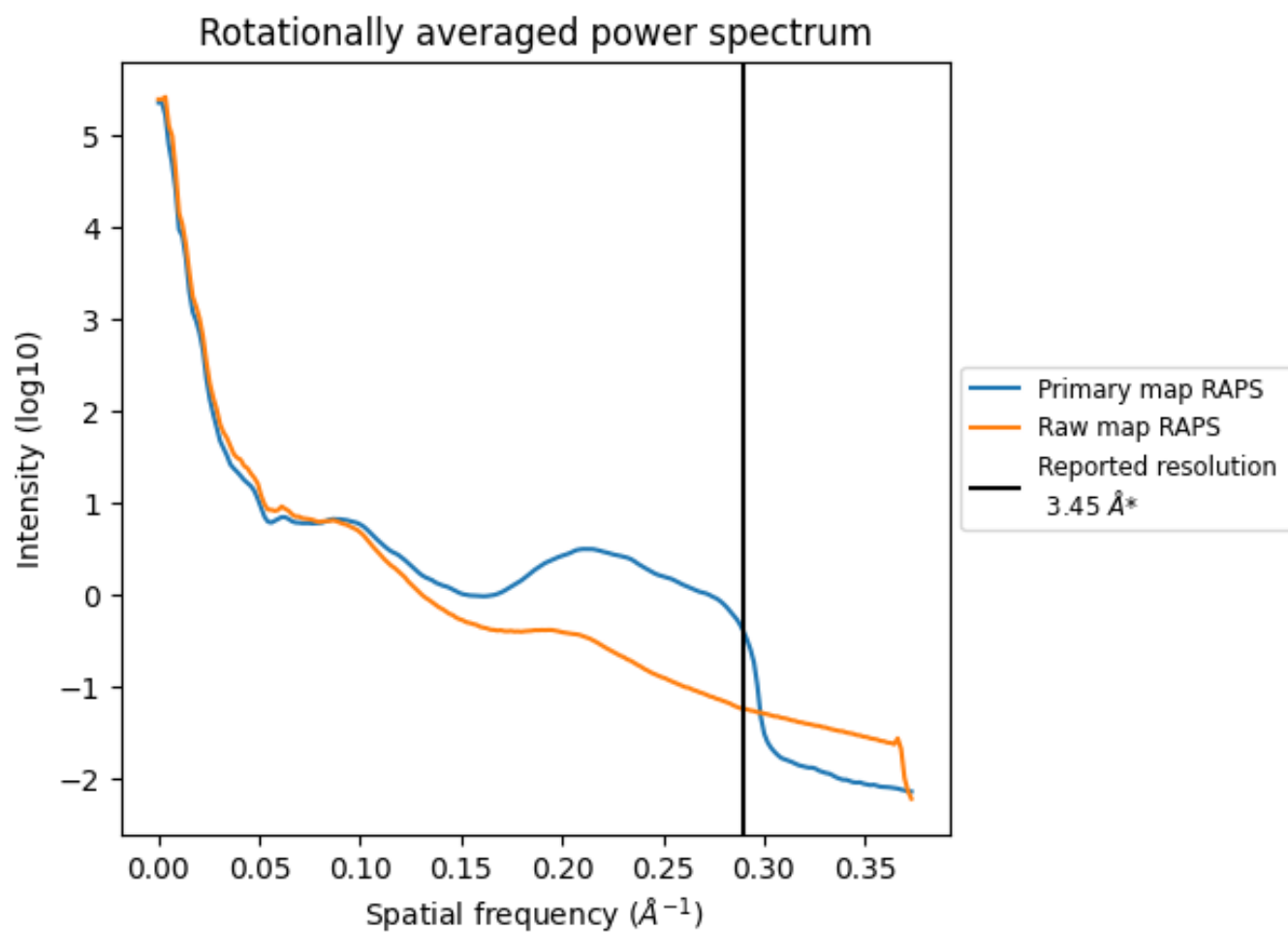
## 7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 66 nm<sup>3</sup>; this corresponds to an approximate mass of 60 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

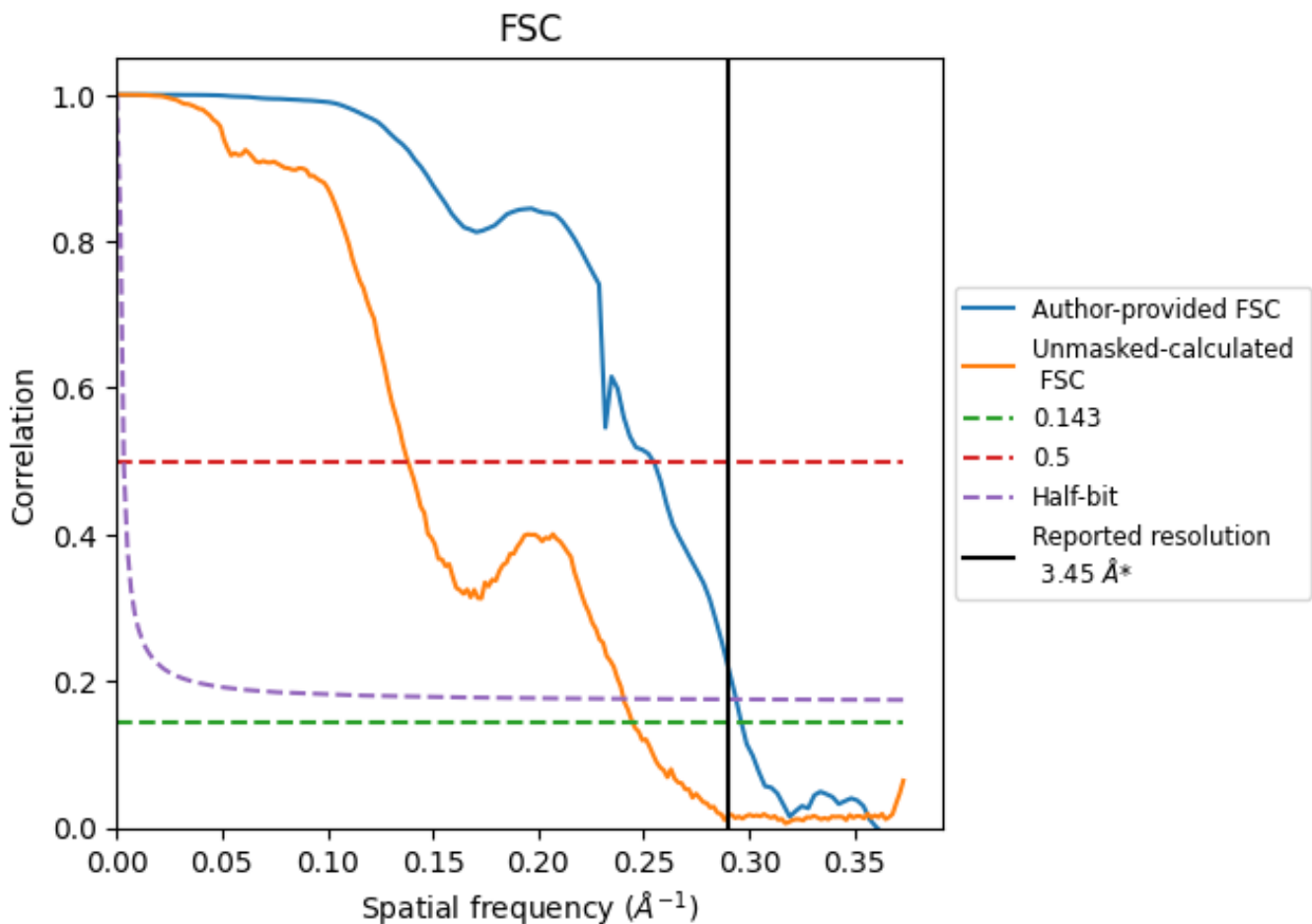


\*Reported resolution corresponds to spatial frequency of 0.290 Å<sup>-1</sup>

## 8 Fourier-Shell correlation [i](#)

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

### 8.1 FSC [i](#)



\*Reported resolution corresponds to spatial frequency of 0.290 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

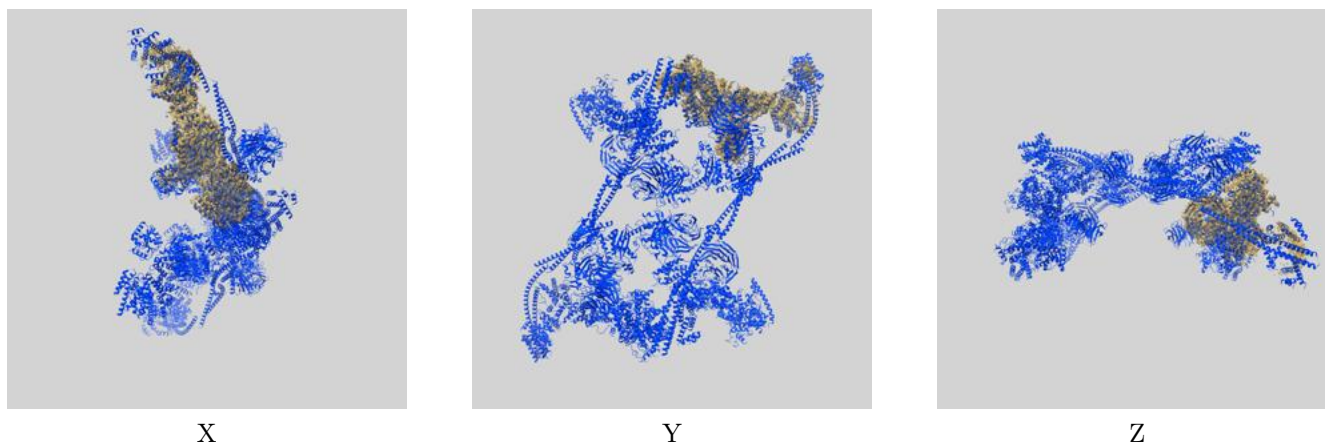
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.45	-	-
Author-provided FSC curve	3.37	3.93	3.41
Unmasked-calculated*	4.08	7.25	4.16

\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 4.08 differs from the reported value 3.45 by more than 10 %

## 9 Map-model fit [i](#)

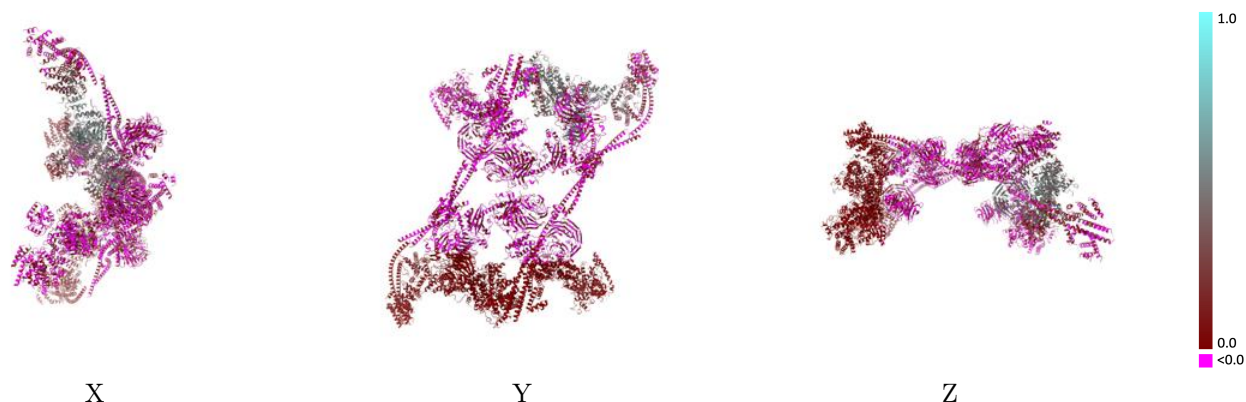
This section contains information regarding the fit between EMDB map EMD-14808 and PDB model 7ZNL. Per-residue inclusion information can be found in section 3 on page 7.

### 9.1 Map-model overlay [i](#)



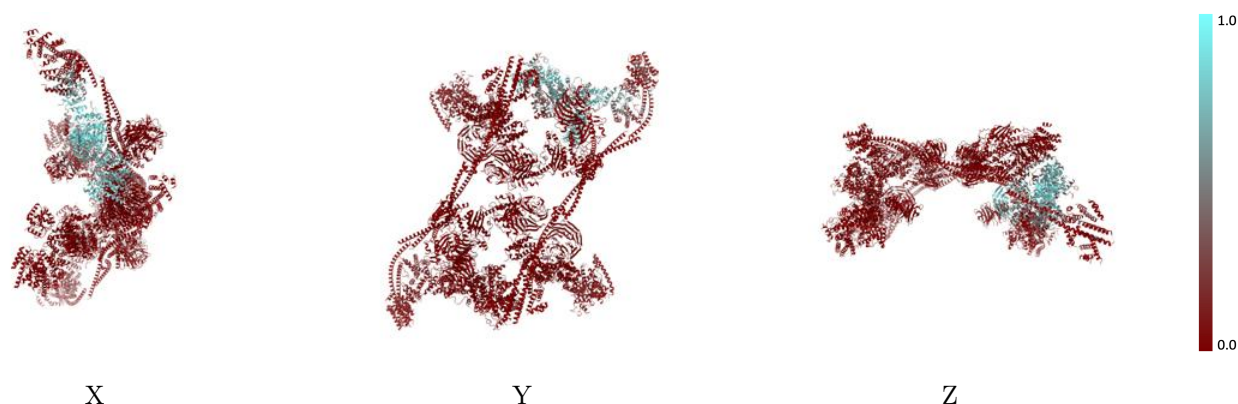
The images above show the 3D surface view of the map at the recommended contour level 0.9 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 9.2 Q-score mapped to coordinate model [i](#)



The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

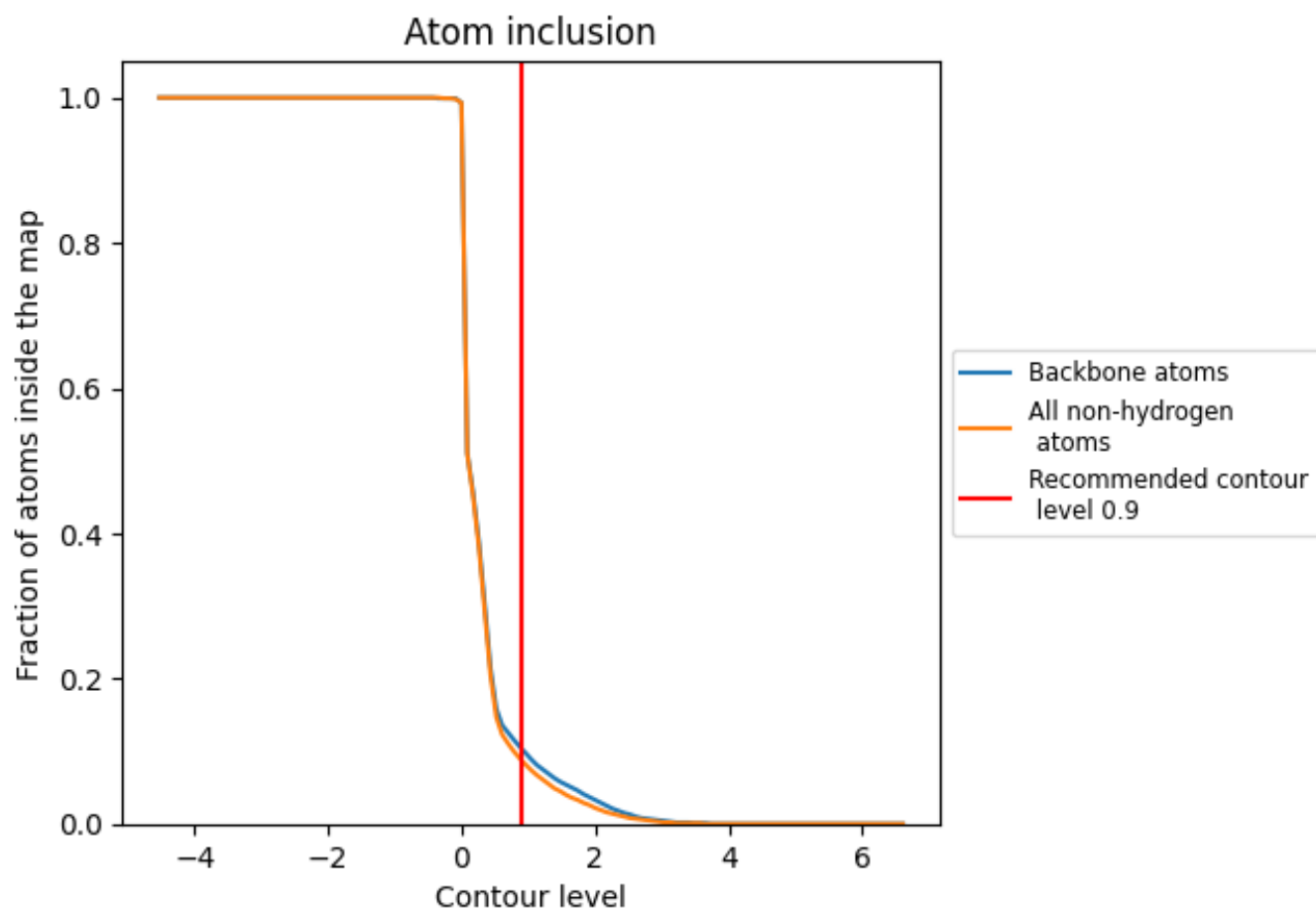
## 9.3 Atom inclusion mapped to coordinate model [i](#)



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 (0.9).





























































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 10% of all backbone atoms, 9% 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 (0.9) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.0870	 0.0690
A	 0.2270	 0.2380
B	 0.6300	 0.4210
C	 0.7970	 0.5090
E	 0.0090	 0.0360
F	 0.0000	 0.0190
G	 0.0750	 0.0590
H	 0.1960	 0.3320
I	 0.0000	 -0.0080
J	 0.0000	 0.0010
K	 0.0000	 -0.0070
M	 0.0000	 0.0170
N	 0.0000	 0.0240
O	 0.0000	 -0.0040
P	 0.0000	 -0.0000
a	 0.0000	 0.0000
b	 0.0000	 0.0000
c	 0.0000	 0.0080
e	 0.0000	 0.0040
f	 0.0000	 0.0200
g	 0.0000	 -0.0010
h	 0.0000	 0.0000
i	 0.0000	 0.0180
j	 0.0000	 0.0040
k	 0.0000	 0.0140
m	 0.0000	 -0.0070
n	 0.0000	 -0.0030
o	 0.0000	 0.0020
p	 0.0000	 -0.0130

