



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 05:35 am GMT

PDB ID : 5B1A  
Title : Bovine heart cytochrome c oxidase in the fully oxidized state at 1.5 angstrom resolution  
Authors : Yano, N.; Muramoto, K.; Shimada, A.; Takemura, S.; Baba, J.; Fujisawa, H.; Mochizuki, M.; Shinzawa-Itoh, K.; Yamashita, E.; Tsukihara, T.; Yoshikawa, S.  
Deposited on : 2015-12-01  
Resolution : 1.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

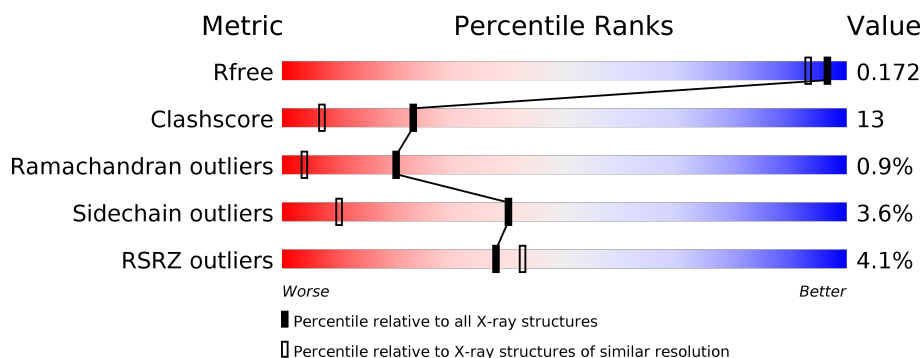
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	2279 (1.50-1.50)
Clashscore	112137	2503 (1.50-1.50)
Ramachandran outliers	110173	2445 (1.50-1.50)
Sidechain outliers	110143	2443 (1.50-1.50)
RSRZ outliers	101464	2305 (1.50-1.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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 electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	514	<div> <div>73%</div> <div>22%</div> <div>.</div> </div>
1	N	514	<div> <div>73%</div> <div>24%</div> <div>.</div> </div>
2	B	227	<div> <div>2%</div> <div>65%</div> <div>29%</div> <div>.</div> </div>
2	O	227	<div> <div>2%</div> <div>69%</div> <div>26%</div> <div>.</div> </div>
3	C	261	<div> <div>72%</div> <div>25%</div> <div>.</div> </div>
3	P	261	<div> <div>%</div> <div>71%</div> <div>22%</div> <div>6%</div> <div>.</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	147	
4	Q	147	
5	E	109	
5	R	109	
6	F	98	
6	S	98	
7	G	85	
7	T	85	
8	H	85	
8	U	85	
9	I	73	
9	V	73	
10	J	59	
10	W	59	
11	K	56	
11	X	56	
12	L	47	
12	Y	47	
13	M	46	
13	Z	46	

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
14	HEA	A	601	X	-	-	-
14	HEA	A	602	X	-	-	-
14	HEA	N	601	X	-	-	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
14	HEA	N	602	X	-	-	-
18	PER	A	606	-	-	-	X
18	PER	N	606	-	-	-	X
19	PGV	A	608	-	-	-	X
19	PGV	C	307	-	-	-	X
19	PGV	Q	201	-	-	-	X
2	FME	B	1	-	-	X	-
20	TGL	B	301	-	-	-	X
20	TGL	D	201	-	-	X	X
20	TGL	L	101	-	-	-	X
20	TGL	N	608	-	-	-	X
20	TGL	Q	202	-	-	-	X
20	TGL	Y	101	-	-	X	X
22	CHD	C	304	-	-	-	X
22	CHD	J	102	-	-	-	X
22	CHD	P	305	-	-	-	X
22	CHD	W	101	-	-	-	X
23	PSC	B	304	-	-	-	X
25	CDL	C	303	-	-	X	X
25	CDL	G	102	-	-	X	X
25	CDL	P	304	-	-	X	X
25	CDL	T	103	-	-	X	X
28	DMU	J	101	-	-	-	X
28	DMU	M	101	-	-	-	X
28	DMU	P	306	-	-	-	X
28	DMU	Z	101	-	-	-	X

## 2 Entry composition

There are 29 unique types of molecules in this entry. The entry contains 35054 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Cytochrome c oxidase subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	514	Total	C	N	O	S	0	18	0
			4168	2778	645	704	41			
1	N	514	Total	C	N	O	S	0	16	0
			4154	2771	643	699	41			

- Molecule 2 is a protein called Cytochrome c oxidase subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	227	Total	C	N	O	S	0	9	0
			1899	1234	292	353	20			
2	O	227	Total	C	N	O	S	0	5	0
			1870	1215	288	347	20			

- Molecule 3 is a protein called Cytochrome c oxidase subunit 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	259	Total	C	N	O	S	0	9	0
			2185	1457	349	363	16			
3	P	259	Total	C	N	O	S	0	9	0
			2185	1457	349	363	16			

- Molecule 4 is a protein called Cytochrome c oxidase subunit 4 isoform 1, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	144	Total	C	N	O	S	0	5	0
			1242	809	206	223	4			
4	Q	144	Total	C	N	O	S	0	3	0
			1224	797	202	221	4			

- Molecule 5 is a protein called Cytochrome c oxidase subunit 5A, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	105	Total	C	N	O	S	0	0	0
			852	544	144	162	2			
5	R	105	Total	C	N	O	S	0	1	0
			863	550	148	163	2			

- Molecule 6 is a protein called Cytochrome c oxidase subunit 5B, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	98	Total	C	N	O	S	0	4	0
			778	481	139	152	6			
6	S	98	Total	C	N	O	S	0	2	0
			763	473	136	148	6			

- Molecule 7 is a protein called Cytochrome c oxidase subunit 6A2, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
7	G	84	Total 686	C 440	N 130	O 114	P 1	S 1	0	1	0
7	T	84	Total 686	C 440	N 130	O 114	P 1	S 1	0	1	0

- Molecule 8 is a protein called Cytochrome c oxidase subunit 6B1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	79	Total	C	N	O	S	0	0	0
			662	417	121	119	5			
8	U	79	Total	C	N	O	S	0	0	0
			662	417	121	119	5			

- Molecule 9 is a protein called Cytochrome c oxidase subunit 6C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	73	Total	C	N	O	S	0	0	0
			601	390	107	100	4			
9	V	73	Total	C	N	O	S	0	0	0
			601	390	107	100	4			

- Molecule 10 is a protein called Cytochrome c oxidase subunit 7A1, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	58	Total	C	N	O	S	0	0	0
			460	297	78	82	3			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	W	58	Total	C	N	O	S	0	1	0
			469	302	79	85	3			

- Molecule 11 is a protein called Cytochrome c oxidase subunit 7B, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	49	Total	C	N	O	S	0	0	0
			384	250	65	67	2			
11	X	49	Total	C	N	O	S	0	1	0
			391	255	66	68	2			

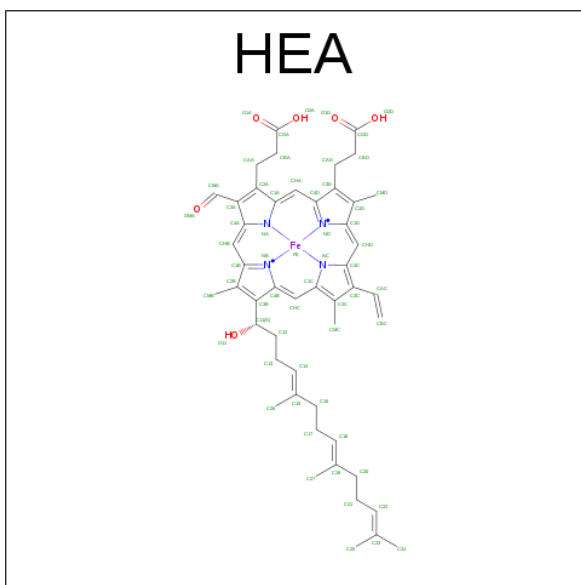
- Molecule 12 is a protein called Cytochrome c oxidase subunit 7C, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	46	Total	C	N	O	S	0	0	0
			380	254	64	60	2			
12	Y	46	Total	C	N	O	S	0	1	0
			388	259	65	61	3			

- Molecule 13 is a protein called Cytochrome c oxidase subunit 8B, mitochondrial.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
13	M	43	Total	C	N	O	0	0	0
			335	223	53	59			
13	Z	43	Total	C	N	O	0	0	0
			335	223	53	59			

- Molecule 14 is HEME-A (three-letter code: HEA) (formula: C<sub>49</sub>H<sub>56</sub>FeN<sub>4</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
14	A	1	Total 60	C 49	Fe 1	N 4	O 6	0	0
14	A	1	Total 60	C 49	Fe 1	N 4	O 6	0	0
14	N	1	Total 60	C 49	Fe 1	N 4	O 6	0	0
14	N	1	Total 60	C 49	Fe 1	N 4	O 6	0	0

- Molecule 15 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	1	Total	Cu	0	0
			1	1		
15	N	1	Total	Cu	0	0
			1	1		

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

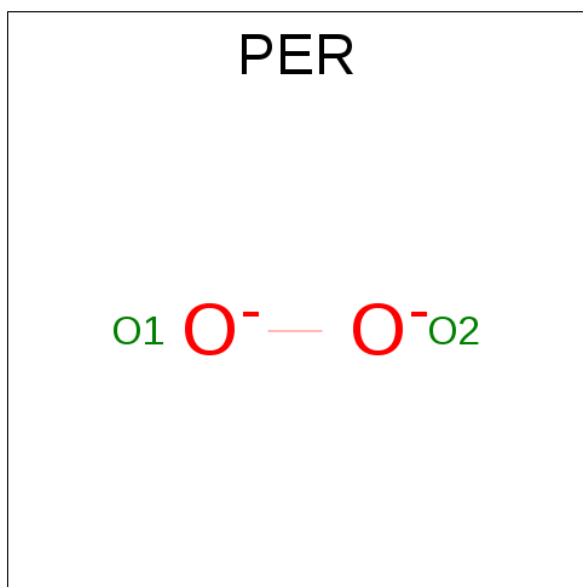
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	1	Total	Mg	0	0
			1	1		
16	N	1	Total	Mg	0	0
			1	1		

- Molecule 17 is SODIUM ION (three-letter code: NA) (formula: Na).



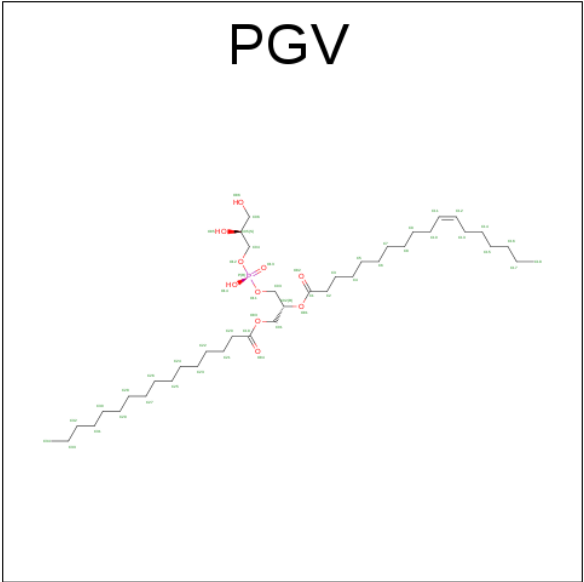
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	A	1	Total	Na	0	0
			1	1		
17	N	1	Total	Na	0	0
			1	1		

- Molecule 18 is PEROXIDE ION (three-letter code: PER) (formula: O<sub>2</sub>).



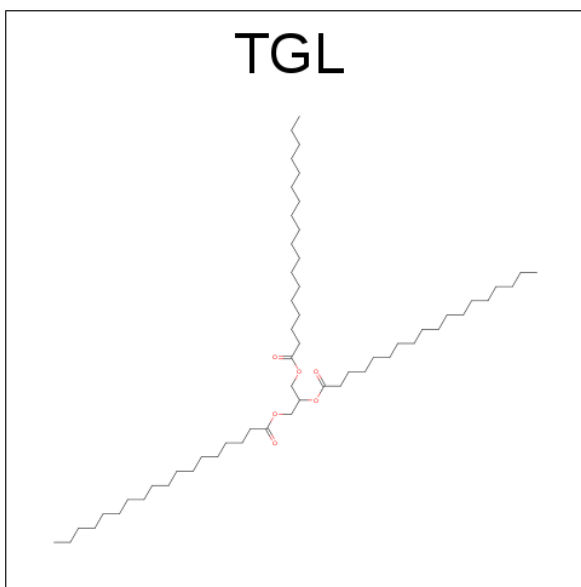
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	A	1	Total	O	0	0
			2	2		
18	N	1	Total	O	0	0
			2	2		

- Molecule 19 is (1R)-2-{{[(2S)-2,3-DIHYDROXYPROPYL]OXY}(HYDROXY)PHOSPHORYL]OXY}-1-[(PALMITOYLOXY)METHYL]ETHYL (11E)-OCTADEC-11-ENOATE (three-letter code: PGV) (formula: C<sub>40</sub>H<sub>77</sub>O<sub>10</sub>P).



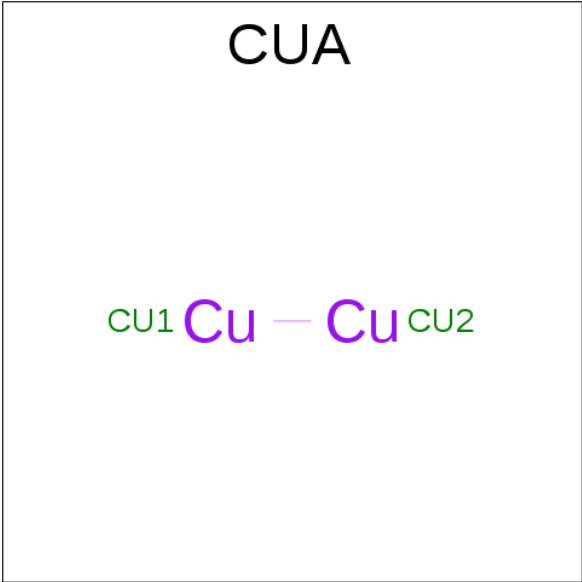
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
19	A	1	Total	C	O	P	0	0
			51	40	10	1		
19	A	1	Total	C	O	P	0	0
			51	40	10	1		
19	C	1	Total	C	O	P	0	0
			51	40	10	1		
19	C	1	Total	C	O	P	0	0
			51	40	10	1		
19	N	1	Total	C	O	P	0	0
			51	40	10	1		
19	P	1	Total	C	O	P	0	0
			51	40	10	1		
19	P	1	Total	C	O	P	0	0
			51	40	10	1		
19	Q	1	Total	C	O	P	0	0
			51	40	10	1		

- Molecule 20 is TRISTEAROYLGLYCEROL (three-letter code: TGL) (formula: C<sub>57</sub>H<sub>110</sub>O<sub>6</sub>).



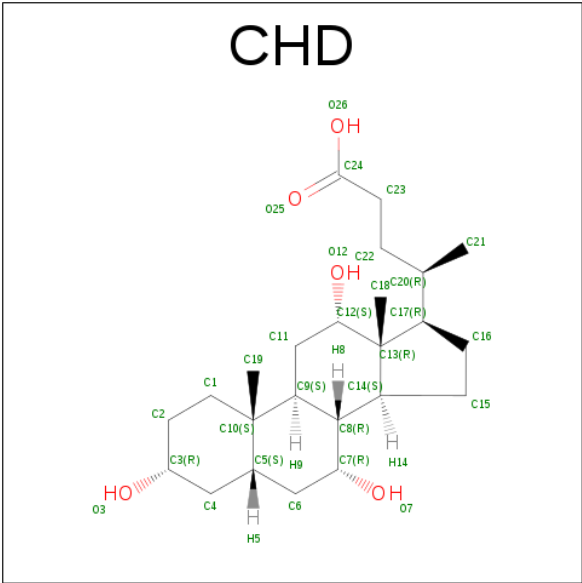
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
20	B	1	Total	C	O	0	0
			63	57	6		
20	D	1	Total	C	O	0	0
			63	57	6		
20	L	1	Total	C	O	0	0
			63	57	6		
20	N	1	Total	C	O	0	0
			63	57	6		
20	Q	1	Total	C	O	0	0
			63	57	6		
20	Y	1	Total	C	O	0	0
			63	57	6		

- Molecule 21 is DINUCLEAR COPPER ION (three-letter code: CUA) (formula: Cu<sub>2</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
21	B	1	Total	Cu	0	0
			2	2		
21	O	1	Total	Cu	0	0
			2	2		

- Molecule 22 is CHOLIC ACID (three-letter code: CHD) (formula: C<sub>24</sub>H<sub>40</sub>O<sub>5</sub>).



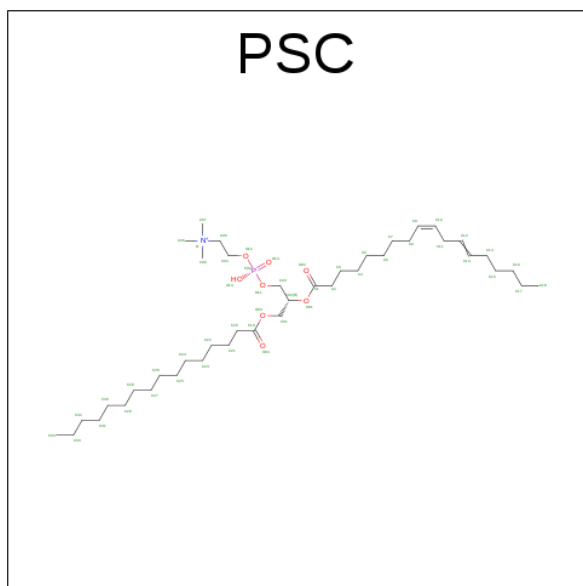
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
22	B	1	Total	C	O	0	0
			29	24	5		
22	C	1	Total	C	O	0	0
			29	24	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
22	C	1	Total	C	O	0	0
			29	24	5		
22	J	1	Total	C	O	0	0
			29	24	5		
22	O	1	Total	C	O	0	0
			29	24	5		
22	P	1	Total	C	O	0	0
			29	24	5		
22	P	1	Total	C	O	0	0
			29	24	5		
22	W	1	Total	C	O	0	0
			29	24	5		

- Molecule 23 is (7R,17E,20E)-4-HYDROXY-N,N,N-TRIMETHYL-9-OXO-7-[(PALMITOYLOXY)METHYL]-3,5,8-TRIOXA-4-PHOSPHAHEXACOSA-17,20-DIEN-1-AMINIUM 4-OXIDE (three-letter code: PSC) (formula: C<sub>42</sub>H<sub>81</sub>NO<sub>8</sub>P).

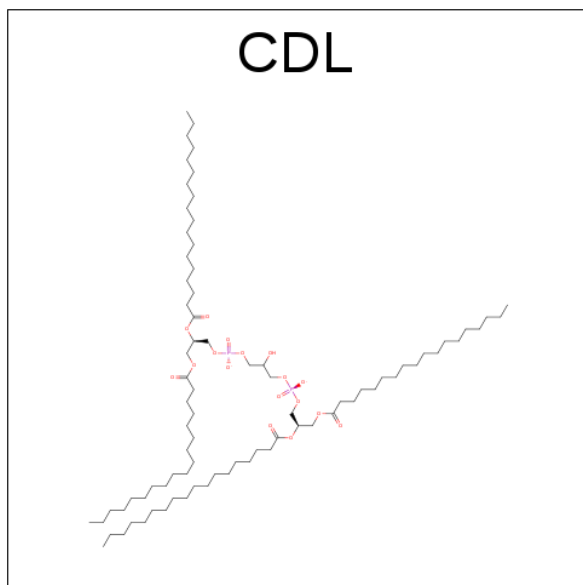


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
23	B	1	Total	C	N	O	P	0	0
			52	42	1	8	1		
23	O	1	Total	C	N	O	P	0	0
			52	42	1	8	1		

- Molecule 24 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

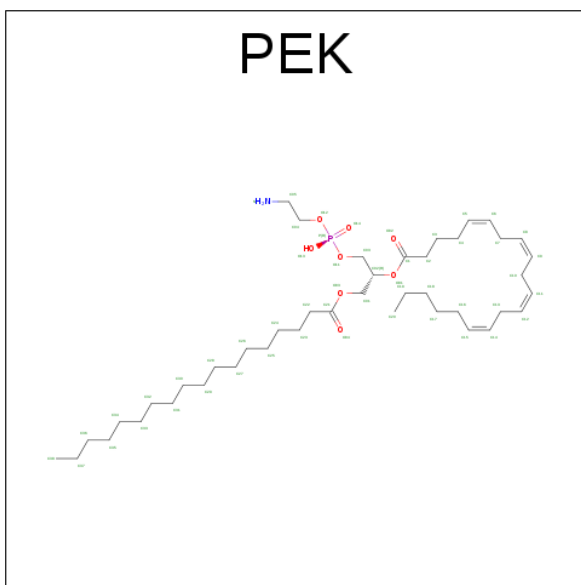
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
24	P	1	Total X 1 1	0	0
24	C	1	Total X 1 1	0	0

- Molecule 25 is CARDIOLIPIN (three-letter code: CDL) (formula:  $C_{81}H_{156}O_{17}P_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
25	C	1	Total C O P 100 81 17 2	0	0
25	G	1	Total C O P 100 81 17 2	0	0
25	P	1	Total C O P 100 81 17 2	0	0
25	T	1	Total C O P 100 81 17 2	0	0

- Molecule 26 is (1S)-2-{[(2-AMINOETHOXY)(HYDROXY)PHOSPHORYL]OXY}-1-[(STEAROYLOXY)METHYL]ETHYL (5E,8E,11E,14E)-ICOSA-5,8,11,14-TETRAENOATE (three-letter code: PEK) (formula:  $C_{43}H_{78}NO_8P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
26	C	1	Total	C	N	O	P	0	0
			53	43	1	8	1		
26	G	1	Total	C	N	O	P	0	0
			53	43	1	8	1		
26	G	1	Total	C	N	O	P	0	0
			53	43	1	8	1		
26	P	1	Total	C	N	O	P	0	0
			53	43	1	8	1		
26	T	1	Total	C	N	O	P	0	0
			53	43	1	8	1		
26	T	1	Total	C	N	O	P	0	0
			53	43	1	8	1		

- Molecule 27 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
27	S	1	Total	Zn	0	0
			1	1		
27	F	1	Total	Zn	0	0
			1	1		

- Molecule 28 is DECYL-BETA-D-MALTOPYRANOSIDE (three-letter code: DMU) (formula: C<sub>22</sub>H<sub>42</sub>O<sub>11</sub>).



- Molecule 29 is water.

*Continued on next page...*



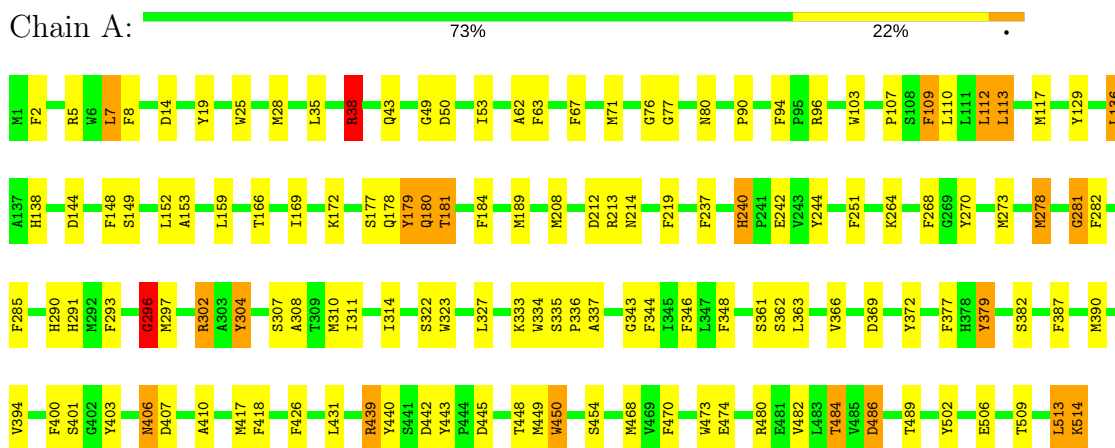
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
29	I	88	Total 88	O 88	0	0
29	J	63	Total 63	O 63	0	0
29	K	69	Total 69	O 69	0	0
29	L	48	Total 48	O 48	0	0
29	M	47	Total 47	O 47	0	0
29	N	290	Total 290	O 290	0	0
29	O	242	Total 243	O 243	0	1
29	P	173	Total 173	O 173	0	0
29	Q	164	Total 164	O 164	0	0
29	R	151	Total 151	O 151	0	0
29	S	186	Total 186	O 186	0	0
29	T	94	Total 94	O 94	0	0
29	U	110	Total 110	O 110	0	0
29	V	71	Total 71	O 71	0	0
29	W	58	Total 58	O 58	0	0
29	X	57	Total 57	O 57	0	0
29	Y	40	Total 40	O 40	0	0
29	Z	37	Total 37	O 37	0	0

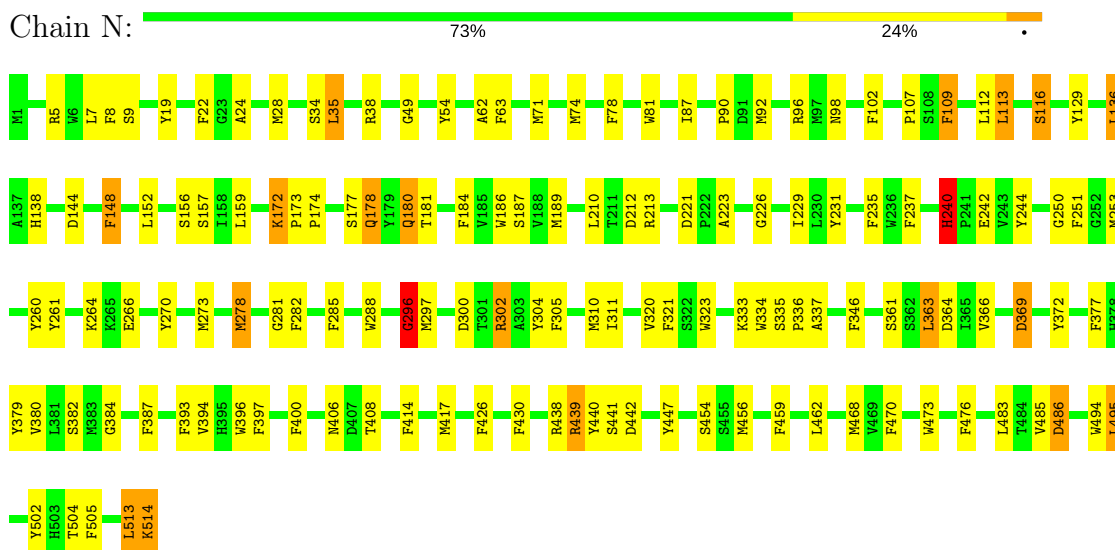
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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 electron 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 dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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: Cytochrome c oxidase subunit 1

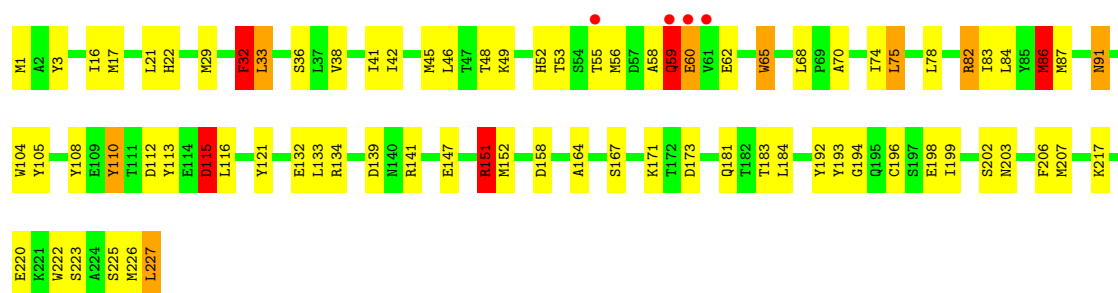


#### • Molecule 1: Cytochrome c oxidase subunit 1

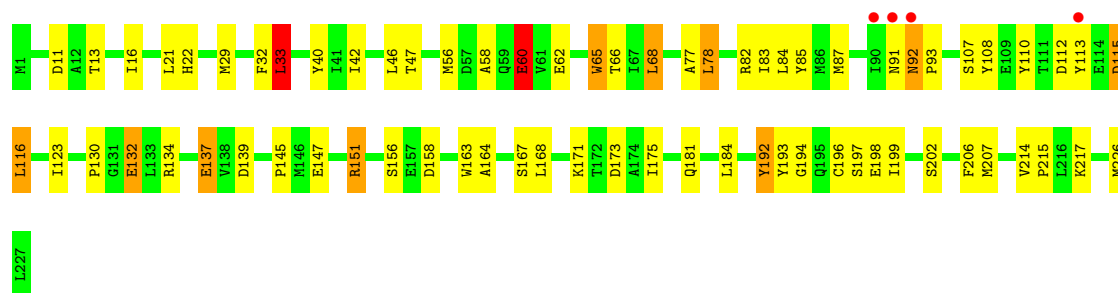


#### • Molecule 2: Cytochrome c oxidase subunit 2

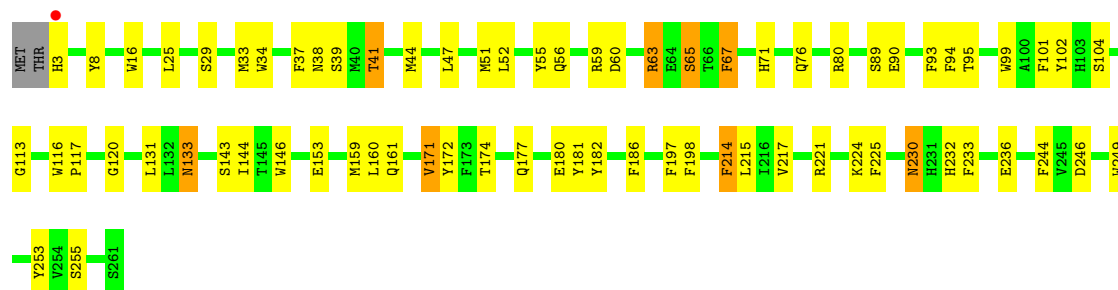




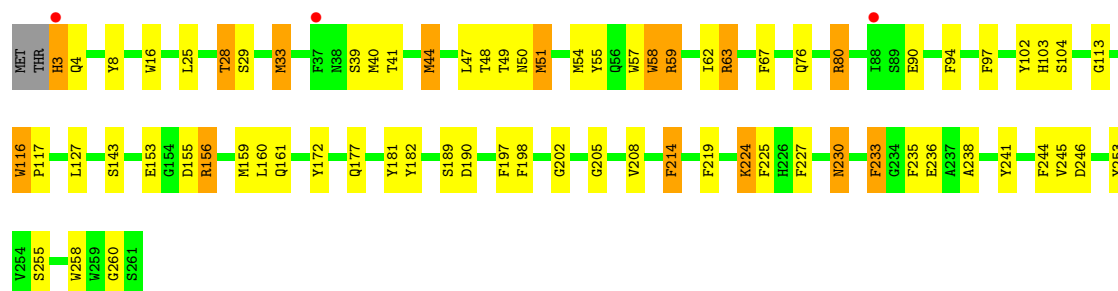
• Molecule 2: Cytochrome c oxidase subunit 2



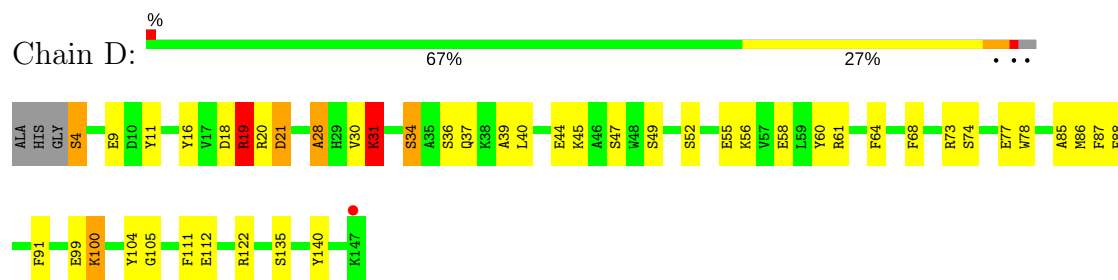
• Molecule 3: Cytochrome c oxidase subunit 3



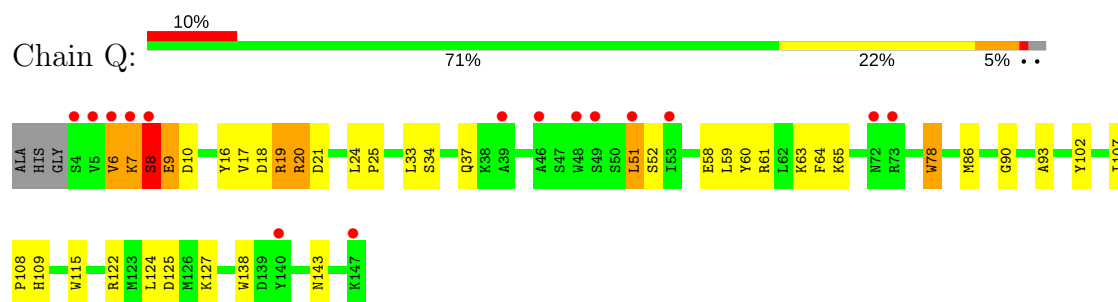
• Molecule 3: Cytochrome c oxidase subunit 3



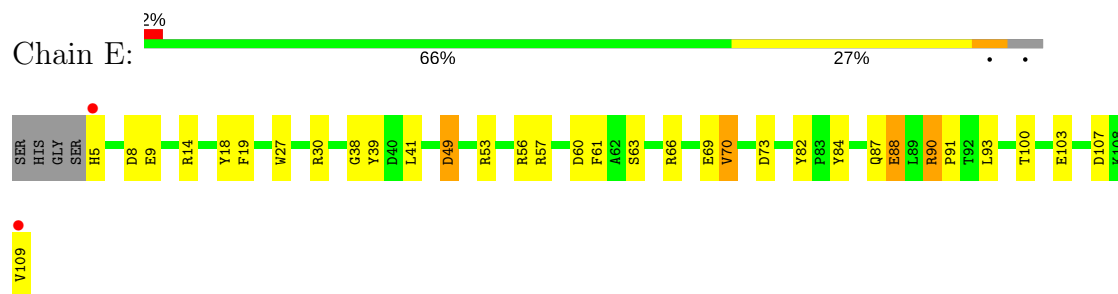
• Molecule 4: Cytochrome c oxidase subunit 4 isoform 1, mitochondrial



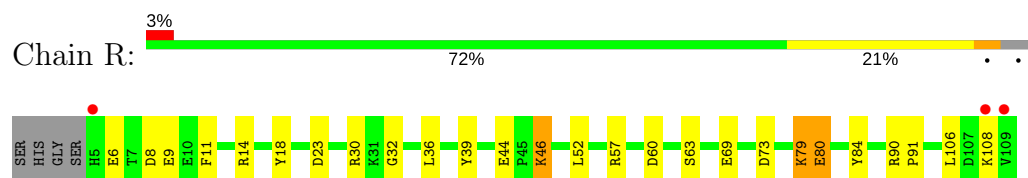
- Molecule 4: Cytochrome c oxidase subunit 4 isoform 1, mitochondrial



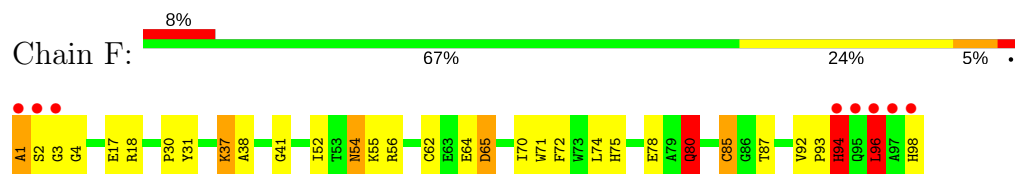
- Molecule 5: Cytochrome c oxidase subunit 5A, mitochondrial



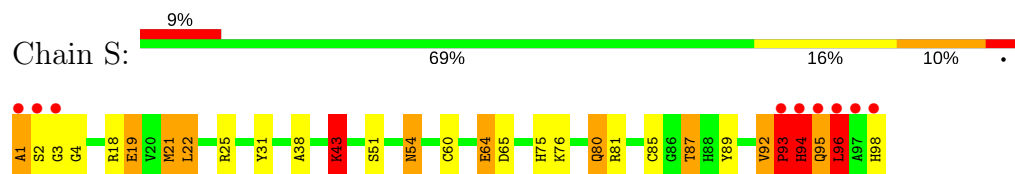
- Molecule 5: Cytochrome c oxidase subunit 5A, mitochondrial



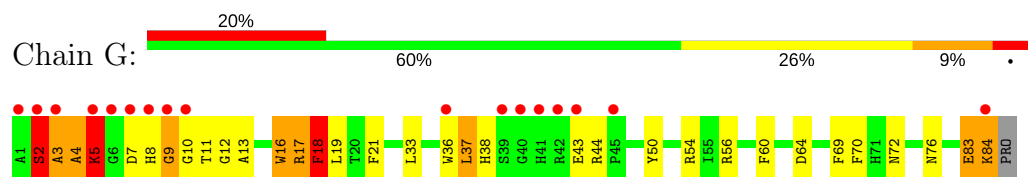
- Molecule 6: Cytochrome c oxidase subunit 5B, mitochondrial



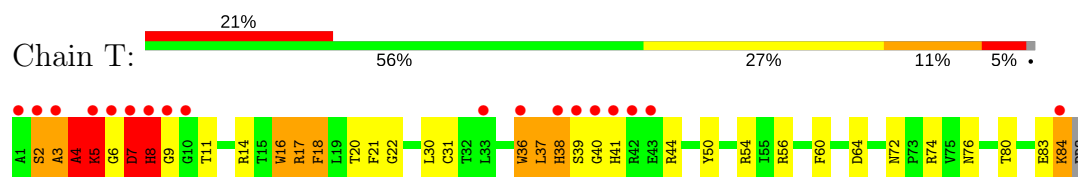
- Molecule 6: Cytochrome c oxidase subunit 5B, mitochondrial



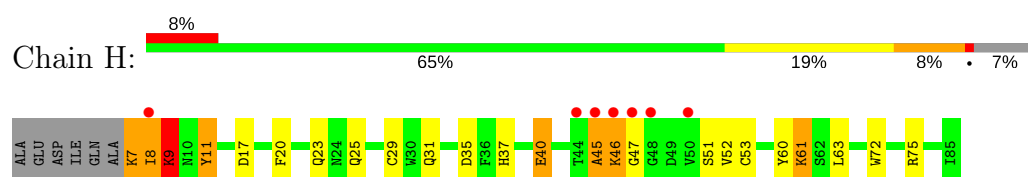
- Molecule 7: Cytochrome c oxidase subunit 6A2, mitochondrial



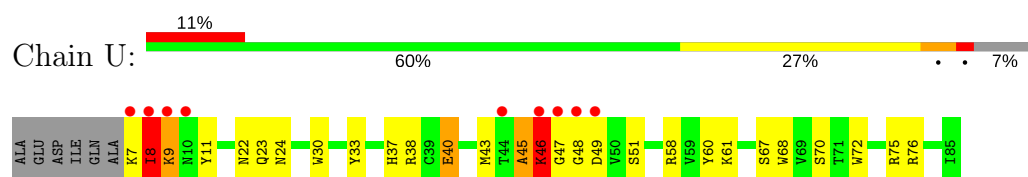
- Molecule 7: Cytochrome c oxidase subunit 6A2, mitochondrial



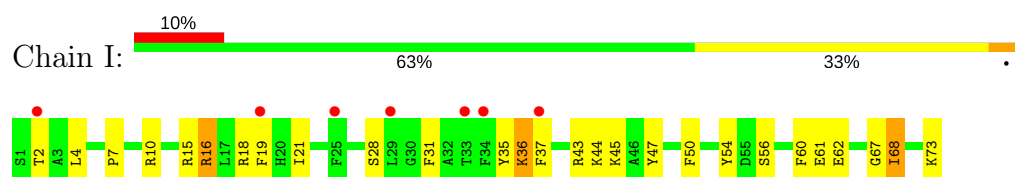
- Molecule 8: Cytochrome c oxidase subunit 6B1



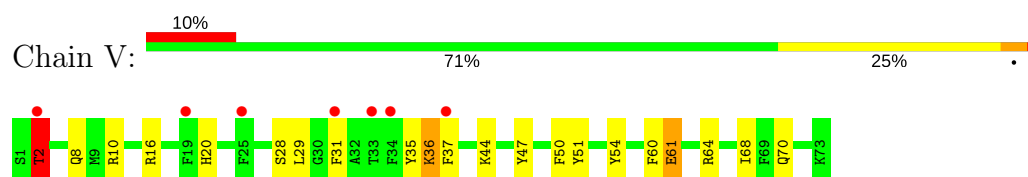
- Molecule 8: Cytochrome c oxidase subunit 6B1



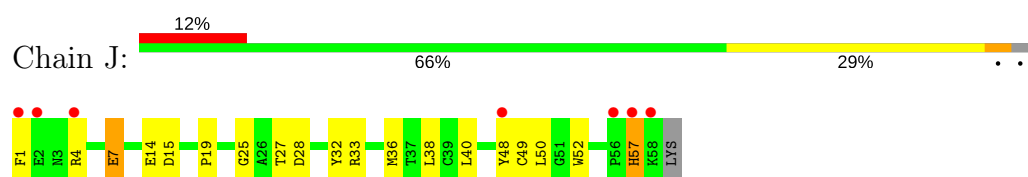
- Molecule 9: Cytochrome c oxidase subunit 6C



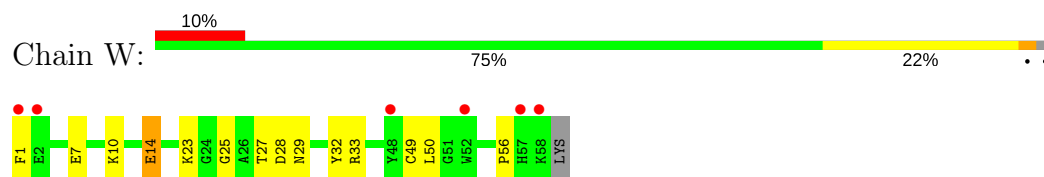
- Molecule 9: Cytochrome c oxidase subunit 6C



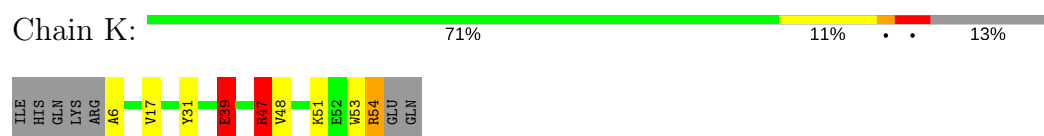
- Molecule 10: Cytochrome c oxidase subunit 7A1, mitochondrial



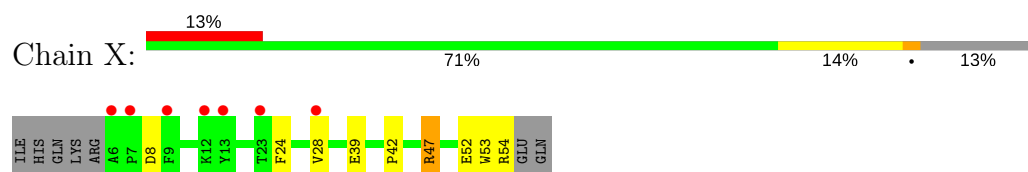
- Molecule 10: Cytochrome c oxidase subunit 7A1, mitochondrial



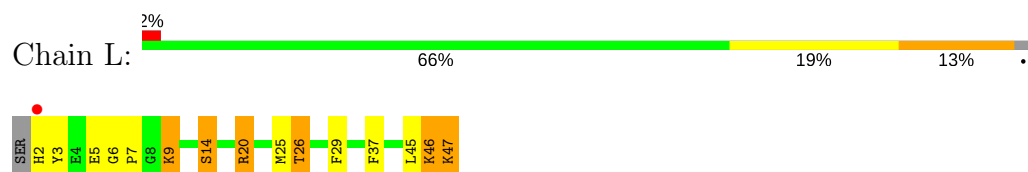
- Molecule 11: Cytochrome c oxidase subunit 7B, mitochondrial



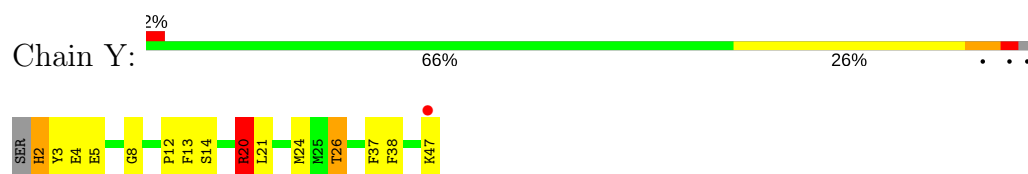
- Molecule 11: Cytochrome c oxidase subunit 7B, mitochondrial



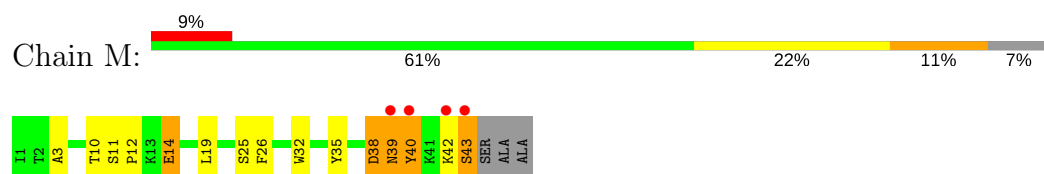
- Molecule 12: Cytochrome c oxidase subunit 7C, mitochondrial



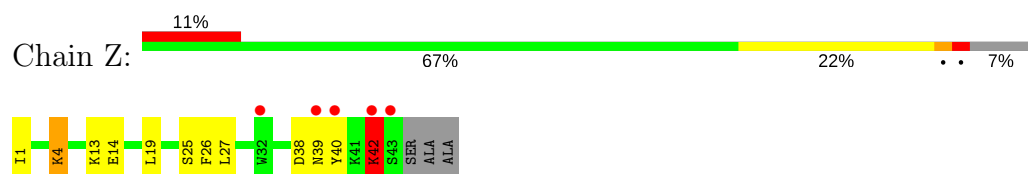
- Molecule 12: Cytochrome c oxidase subunit 7C, mitochondrial



- Molecule 13: Cytochrome c oxidase subunit 8B, mitochondrial



- Molecule 13: Cytochrome c oxidase subunit 8B, mitochondrial



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	181.94Å 204.40Å 177.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.50 89.10 – 1.40	Depositor EDS
% Data completeness (in resolution range)	98.3 (40.00-1.50) 98.2 (89.10-1.40)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.86 (at 1.40Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.149 , 0.172 0.149 , 0.172	Depositor DCC
$R_{free}$ test set	51656 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	20.6	Xtriage
Anisotropy	0.642	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 66.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.001 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	35054	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.68% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CHD, HEA, SAC, CDL, PSC, PEK, MG, PER, PGV, TPO, UNX, CUA, NA, FME, TGL, CU, DMU

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	1.84	54/4297 (1.3%)	1.89	89/5864 (1.5%)
1	N	1.85	57/4283 (1.3%)	1.76	84/5845 (1.4%)
2	B	1.94	37/1937 (1.9%)	1.80	32/2637 (1.2%)
2	O	1.82	30/1908 (1.6%)	1.57	23/2597 (0.9%)
3	C	1.91	41/2272 (1.8%)	1.79	46/3102 (1.5%)
3	P	1.91	37/2272 (1.6%)	1.77	52/3102 (1.7%)
4	D	2.05	31/1277 (2.4%)	1.87	40/1720 (2.3%)
4	Q	1.67	17/1259 (1.4%)	1.88	23/1698 (1.4%)
5	E	2.01	24/871 (2.8%)	2.45	41/1182 (3.5%)
5	R	1.89	18/882 (2.0%)	1.60	14/1196 (1.2%)
6	F	1.96	16/795 (2.0%)	1.65	7/1079 (0.6%)
6	S	1.89	17/780 (2.2%)	1.69	14/1058 (1.3%)
7	G	2.03	15/702 (2.1%)	1.87	19/953 (2.0%)
7	T	1.95	14/702 (2.0%)	1.59	9/953 (0.9%)
8	H	1.77	7/682 (1.0%)	1.52	9/921 (1.0%)
8	U	1.74	10/682 (1.5%)	1.36	5/921 (0.5%)
9	I	1.96	13/605 (2.1%)	1.73	13/802 (1.6%)
9	V	1.70	7/605 (1.2%)	1.85	9/802 (1.1%)
10	J	1.80	6/471 (1.3%)	1.54	4/636 (0.6%)
10	W	1.65	4/480 (0.8%)	1.36	2/648 (0.3%)
11	K	2.09	14/398 (3.5%)	1.91	7/546 (1.3%)
11	X	1.63	7/405 (1.7%)	1.47	5/556 (0.9%)
12	L	2.01	6/393 (1.5%)	1.76	11/526 (2.1%)
12	Y	1.97	11/401 (2.7%)	1.52	3/536 (0.6%)
13	M	1.80	6/345 (1.7%)	1.68	5/470 (1.1%)
13	Z	1.70	4/345 (1.2%)	1.43	3/470 (0.6%)
All	All	1.87	503/30049 (1.7%)	1.77	569/40820 (1.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected



by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	N	0	3
2	B	0	3
3	C	0	1
4	Q	0	1
5	E	0	2
6	F	0	1
6	S	0	2
7	G	0	1
7	T	0	1
10	J	0	1
11	K	0	1
13	M	0	1
All	All	0	22

The worst 5 of 503 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	O	65	TRP	CB-CG	-17.96	1.18	1.50
2	B	65	TRP	CB-CG	-17.06	1.19	1.50
4	D	58	GLU	CD-OE1	15.90	1.43	1.25
11	K	47	ARG	CZ-NH2	14.60	1.52	1.33
7	T	36	TRP	CB-CG	13.34	1.74	1.50

The worst 5 of 569 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	90	ARG	NE-CZ-NH1	35.49	138.05	120.30
4	Q	20	ARG	NE-CZ-NH1	32.93	136.77	120.30
4	Q	20	ARG	NE-CZ-NH2	-31.64	104.48	120.30
5	E	90	ARG	NE-CZ-NH2	-24.93	107.83	120.30
9	V	10	ARG	NE-CZ-NH2	-21.32	109.64	120.30

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	296	GLY	Mainchain
1	A	304	TYR	Sidechain
1	A	379	TYR	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	A	38	ARG	Sidechain
2	B	68	LEU	Mainchain

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4168	0	4137	74	0
1	N	4154	0	4129	81	0
2	B	1899	0	1898	65	0
2	O	1870	0	1868	40	0
3	C	2185	0	2097	38	0
3	P	2185	0	2097	51	0
4	D	1242	0	1235	19	0
4	Q	1224	0	1211	25	0
5	E	852	0	845	1	0
5	R	863	0	857	8	2
6	F	778	0	754	27	0
6	S	763	0	742	42	0
7	G	686	0	651	42	0
7	T	686	0	651	45	0
8	H	662	0	623	24	0
8	U	662	0	623	15	0
9	I	601	0	613	17	2
9	V	601	0	613	13	0
10	J	460	0	459	13	0
10	W	469	0	464	9	0
11	K	384	0	366	6	0
11	X	391	0	374	4	0
12	L	380	0	380	19	0
12	Y	388	0	388	29	0
13	M	335	0	352	7	0
13	Z	335	0	352	5	0
14	A	120	0	107	10	0
14	N	120	0	107	11	0
15	A	1	0	0	0	0
15	N	1	0	0	0	0
16	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	N	1	0	0	0	0
17	A	1	0	0	0	0
17	N	1	0	0	0	0
18	A	2	0	0	1	0
18	N	2	0	0	1	0
19	A	102	0	152	11	0
19	C	102	0	152	8	0
19	N	51	0	76	1	0
19	P	102	0	152	9	0
19	Q	51	0	76	11	0
20	B	63	0	109	3	0
20	D	63	0	106	21	0
20	L	63	0	110	19	0
20	N	63	0	110	6	0
20	Q	63	0	110	15	0
20	Y	63	0	110	23	0
21	B	2	0	0	0	0
21	O	2	0	0	0	0
22	B	29	0	39	0	0
22	C	58	0	77	3	0
22	J	29	0	37	6	0
22	O	29	0	39	0	0
22	P	58	0	77	7	0
22	W	29	0	38	6	0
23	B	52	0	80	13	0
23	O	52	0	80	19	0
24	C	1	0	0	0	0
24	P	1	0	0	0	0
25	C	100	0	154	26	0
25	G	100	0	156	37	0
25	P	100	0	156	24	0
25	T	100	0	156	28	0
26	C	53	0	77	13	0
26	G	106	0	154	16	0
26	P	53	0	77	16	0
26	T	106	0	154	18	0
27	F	1	0	0	0	0
27	S	1	0	0	0	0
28	J	33	0	42	5	0
28	M	33	0	42	0	0
28	P	33	0	42	10	0
28	Z	33	0	42	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
29	A	297	0	0	17	0
29	B	274	0	0	13	0
29	C	176	0	0	8	0
29	D	266	0	0	5	0
29	E	178	0	0	2	0
29	F	199	0	0	9	0
29	G	100	0	0	7	0
29	H	122	0	0	8	0
29	I	88	0	0	2	0
29	J	63	0	0	2	0
29	K	69	0	0	4	0
29	L	48	0	0	2	0
29	M	47	0	0	0	0
29	N	290	0	0	12	0
29	O	243	0	0	4	0
29	P	173	0	0	8	0
29	Q	164	0	0	8	0
29	R	151	0	0	4	0
29	S	186	0	0	9	0
29	T	94	0	0	3	0
29	U	110	0	0	5	0
29	V	71	0	0	2	0
29	W	58	0	0	1	0
29	X	57	0	0	1	0
29	Y	40	0	0	3	0
29	Z	37	0	0	0	0
All	All	35054	0	31975	816	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 13.

The worst 5 of 816 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)
19:C:302:PGV:C21	19:C:302:PGV:C22	1.77	1.62
26:T:101:PEK:C2	26:T:101:PEK:C3	1.76	1.56
2:B:1:FME:CN	2:B:1:FME:N	1.70	1.53
3:P:224:LYS:NZ	3:P:224:LYS:CE	1.77	1.47
20:D:201:TGL:OG2	20:D:201:TGL:CB1	1.63	1.47

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:I:2:THR:CB	5:R:80:GLU:OE1[3_647]	2.00	0.20
9:I:2:THR:CG2	5:R:80:GLU:OE1[3_647]	2.19	0.01

## 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 X-ray entries followed by that with respect to entries of similar resolution.

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	530/514 (103%)	515 (97%)	15 (3%)	0	100	100
1	N	528/514 (103%)	511 (97%)	17 (3%)	0	100	100
2	B	234/227 (103%)	230 (98%)	4 (2%)	0	100	100
2	O	230/227 (101%)	225 (98%)	4 (2%)	1 (0%)	38	14
3	C	266/261 (102%)	261 (98%)	4 (2%)	1 (0%)	38	14
3	P	266/261 (102%)	261 (98%)	5 (2%)	0	100	100
4	D	147/147 (100%)	143 (97%)	4 (3%)	0	100	100
4	Q	145/147 (99%)	139 (96%)	5 (3%)	1 (1%)	25	6
5	E	103/109 (94%)	102 (99%)	1 (1%)	0	100	100
5	R	104/109 (95%)	104 (100%)	0	0	100	100
6	F	100/98 (102%)	95 (95%)	3 (3%)	2 (2%)	9	1
6	S	98/98 (100%)	92 (94%)	2 (2%)	4 (4%)	3	0
7	G	82/85 (96%)	68 (83%)	8 (10%)	6 (7%)	1	0
7	T	82/85 (96%)	71 (87%)	5 (6%)	6 (7%)	1	0
8	H	77/85 (91%)	68 (88%)	6 (8%)	3 (4%)	3	0
8	U	77/85 (91%)	68 (88%)	5 (6%)	4 (5%)	2	0
9	I	71/73 (97%)	71 (100%)	0	0	100	100
9	V	71/73 (97%)	69 (97%)	1 (1%)	1 (1%)	13	2
10	J	56/59 (95%)	56 (100%)	0	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
10	W	57/59 (97%)	57 (100%)	0	0	100	100
11	K	47/56 (84%)	45 (96%)	2 (4%)	0	100	100
11	X	48/56 (86%)	46 (96%)	2 (4%)	0	100	100
12	L	44/47 (94%)	42 (96%)	2 (4%)	0	100	100
12	Y	45/47 (96%)	43 (96%)	2 (4%)	0	100	100
13	M	41/46 (89%)	39 (95%)	2 (5%)	0	100	100
13	Z	41/46 (89%)	39 (95%)	1 (2%)	1 (2%)	7	0
All	All	3590/3614 (99%)	3460 (96%)	100 (3%)	30 (1%)	20	5

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	F	94	HIS
6	F	96	LEU
7	G	4	ALA
7	G	5	LYS
7	G	8	HIS

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	444/426 (104%)	438 (99%)	6 (1%)	71	45
1	N	442/426 (104%)	432 (98%)	10 (2%)	56	23
2	B	219/210 (104%)	208 (95%)	11 (5%)	28	4
2	O	215/210 (102%)	206 (96%)	9 (4%)	34	6
3	C	233/226 (103%)	230 (99%)	3 (1%)	73	48
3	P	233/226 (103%)	230 (99%)	3 (1%)	73	48
4	D	133/129 (103%)	128 (96%)	5 (4%)	38	9
4	Q	131/129 (102%)	124 (95%)	7 (5%)	26	4
5	E	92/95 (97%)	90 (98%)	2 (2%)	57	24

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	R	93/95 (98%)	91 (98%)	2 (2%)	57	24
6	F	85/81 (105%)	79 (93%)	6 (7%)	17	1
6	S	83/81 (102%)	75 (90%)	8 (10%)	10	0
7	G	68/68 (100%)	63 (93%)	5 (7%)	16	1
7	T	68/68 (100%)	61 (90%)	7 (10%)	8	0
8	H	71/75 (95%)	65 (92%)	6 (8%)	12	1
8	U	71/75 (95%)	67 (94%)	4 (6%)	25	3
9	I	57/57 (100%)	55 (96%)	2 (4%)	41	11
9	V	57/57 (100%)	52 (91%)	5 (9%)	12	1
10	J	49/50 (98%)	49 (100%)	0	100	100
10	W	50/50 (100%)	48 (96%)	2 (4%)	36	7
11	K	39/46 (85%)	38 (97%)	1 (3%)	51	18
11	X	40/46 (87%)	40 (100%)	0	100	100
12	L	39/40 (98%)	38 (97%)	1 (3%)	51	18
12	Y	40/40 (100%)	38 (95%)	2 (5%)	28	4
13	M	37/38 (97%)	33 (89%)	4 (11%)	7	0
13	Z	37/38 (97%)	33 (89%)	4 (11%)	7	0
All	All	3126/3082 (101%)	3011 (96%)	115 (4%)	40	9

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	N	109	PHE
2	O	65	TRP
9	V	70	GLN
1	N	138	HIS
1	N	363	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
10	J	29	ASN
2	O	91	ASN
9	V	20	HIS
10	J	57	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	N	180	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	FME	A	1	1	9,9,10	2.57	3 (33%)	7,9,11	1.74	1 (14%)
2	FME	B	1	2	9,9,10	5.28	6 (66%)	7,9,11	8.40	3 (42%)
7	TPO	G	11	7	9,10,11	2.03	3 (33%)	10,14,16	2.49	3 (30%)
9	SAC	I	1	9	8,8,9	2.45	3 (37%)	6,9,11	1.54	1 (16%)
1	FME	N	1	1	9,9,10	1.43	2 (22%)	7,9,11	1.97	2 (28%)
2	FME	O	1	2	9,9,10	1.77	4 (44%)	7,9,11	2.05	3 (42%)
7	TPO	T	11	7	9,10,11	2.20	3 (33%)	10,14,16	1.39	2 (20%)
9	SAC	V	1	9	8,8,9	2.09	2 (25%)	6,9,11	2.33	3 (50%)

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
1	FME	A	1	1	-	0/6/9/11	0/0/0/0
2	FME	B	1	2	-	1/6/9/11	0/0/0/0
7	TPO	G	11	7	-	0/8/11/13	0/0/0/0
9	SAC	I	1	9	-	0/6/8/10	0/0/0/0
1	FME	N	1	1	-	0/6/9/11	0/0/0/0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	FME	O	1	2	-	0/6/9/11	0/0/0/0
7	TPO	T	11	7	-	0/8/11/13	0/0/0/0
9	SAC	V	1	9	-	0/6/8/10	0/0/0/0

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	FME	O1-CN	-8.60	0.95	1.22
2	B	1	FME	CG-SD	-4.02	1.59	1.81
2	O	1	FME	CG-SD	-2.90	1.65	1.81
1	N	1	FME	CB-CG	2.09	1.59	1.51
7	T	11	TPO	P-OG1	2.10	1.63	1.59

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1	FME	O1-CN-N	-16.25	79.86	125.20
2	B	1	FME	CA-N-CN	-13.95	101.37	122.82
2	B	1	FME	CG-CB-CA	-5.41	97.31	112.97
7	G	11	TPO	C-CA-N	-4.49	100.81	109.86
2	O	1	FME	CG-CB-CA	-4.15	100.96	112.97

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	FME	O1-CN-N-CA

There are no ring outliers.

3 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1	FME	7	0
7	G	11	TPO	4	0
7	T	11	TPO	3	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 56 ligands modelled in this entry, 8 are monoatomic and 2 are unknown - leaving 46 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
14	HEA	A	601	1	44,67,67	2.52	16 (36%)	37,103,103	2.86	16 (43%)
14	HEA	A	602	1,18	44,67,67	1.78	9 (20%)	37,103,103	2.26	14 (37%)
18	PER	A	606	15,14	0,1,1	0.00	-	0,0,0	0.00	-
19	PGV	A	607	-	50,50,50	1.50	9 (18%)	51,56,56	1.57	11 (21%)
19	PGV	A	608	-	50,50,50	1.94	6 (12%)	51,56,56	2.53	15 (29%)
20	TGL	B	301	-	62,62,62	1.71	8 (12%)	65,65,65	2.98	21 (32%)
21	CUA	B	302	2	0,1,1	0.00	-	0,0,0	0.00	-
22	CHD	B	303	-	29,32,32	2.73	17 (58%)	47,51,51	2.71	22 (46%)
23	PSC	B	304	-	51,51,51	1.32	3 (5%)	56,59,59	1.81	14 (25%)
19	PGV	C	302	-	50,50,50	1.40	5 (10%)	51,56,56	2.07	5 (9%)
25	CDL	C	303	-	99,99,99	1.78	20 (20%)	101,111,111	2.35	31 (30%)
22	CHD	C	304	-	29,32,32	1.47	5 (17%)	47,51,51	3.81	27 (57%)
22	CHD	C	305	-	29,32,32	2.42	13 (44%)	47,51,51	2.74	19 (40%)
26	PEK	C	306	-	52,52,52	1.60	5 (9%)	54,57,57	2.26	19 (35%)
19	PGV	C	307	-	50,50,50	1.27	4 (8%)	51,56,56	1.67	7 (13%)
20	TGL	D	201	-	62,62,62	2.74	11 (17%)	65,65,65	3.43	25 (38%)
26	PEK	G	101	-	52,52,52	1.21	4 (7%)	54,57,57	1.75	13 (24%)
25	CDL	G	102	-	99,99,99	1.47	14 (14%)	101,111,111	1.89	23 (22%)
26	PEK	G	103	-	52,52,52	1.28	3 (5%)	54,57,57	1.86	12 (22%)
28	DMU	J	101	-	34,34,34	0.99	1 (2%)	45,45,45	1.52	10 (22%)
22	CHD	J	102	-	29,32,32	1.84	8 (27%)	47,51,51	4.74	32 (68%)
20	TGL	L	101	-	62,62,62	2.08	13 (20%)	65,65,65	2.90	26 (40%)
28	DMU	M	101	-	34,34,34	1.52	5 (14%)	45,45,45	2.28	17 (37%)
14	HEA	N	601	1	44,67,67	2.33	15 (34%)	37,103,103	3.13	12 (32%)
14	HEA	N	602	1,18	44,67,67	1.69	11 (25%)	37,103,103	2.63	15 (40%)
18	PER	N	606	15,14	0,1,1	0.00	-	0,0,0	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
19	PGV	N	607	-	50,50,50	1.48	7 (14%)	51,56,56	1.53	10 (19%)
20	TGL	N	608	-	62,62,62	1.43	8 (12%)	65,65,65	2.27	15 (23%)
21	CUA	O	301	2	0,1,1	0.00	-	0,0,0	0.00	-
22	CHD	O	302	-	29,32,32	2.45	10 (34%)	47,51,51	2.50	21 (44%)
23	PSC	O	303	-	51,51,51	1.44	3 (5%)	56,59,59	1.76	10 (17%)
19	PGV	P	301	-	50,50,50	1.09	2 (4%)	51,56,56	1.74	10 (19%)
19	PGV	P	303	-	50,50,50	1.06	3 (6%)	51,56,56	1.57	13 (25%)
25	CDL	P	304	-	99,99,99	2.01	24 (24%)	101,111,111	2.22	34 (33%)
22	CHD	P	305	-	29,32,32	1.39	5 (17%)	47,51,51	3.60	28 (59%)
28	DMU	P	306	-	34,34,34	0.86	1 (2%)	45,45,45	1.90	11 (24%)
22	CHD	P	307	-	29,32,32	1.88	9 (31%)	47,51,51	2.50	18 (38%)
26	PEK	P	308	-	52,52,52	1.66	6 (11%)	54,57,57	1.97	13 (24%)
19	PGV	Q	201	-	50,50,50	1.30	3 (6%)	51,56,56	1.81	11 (21%)
20	TGL	Q	202	-	62,62,62	2.33	10 (16%)	65,65,65	2.81	14 (21%)
26	PEK	T	101	-	52,52,52	1.61	5 (9%)	54,57,57	2.32	11 (20%)
26	PEK	T	102	-	52,52,52	1.30	2 (3%)	54,57,57	1.66	8 (14%)
25	CDL	T	103	-	99,99,99	1.50	12 (12%)	101,111,111	1.75	18 (17%)
22	CHD	W	101	-	29,32,32	1.96	10 (34%)	47,51,51	4.67	29 (61%)
20	TGL	Y	101	-	62,62,62	2.03	12 (19%)	65,65,65	3.14	27 (41%)
28	DMU	Z	101	-	34,34,34	1.28	5 (14%)	45,45,45	1.92	13 (28%)

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
14	HEA	A	601	1	3/3/7/16	0/24/76/76	0/0/8/8
14	HEA	A	602	1,18	3/3/7/16	0/24/76/76	0/0/8/8
18	PER	A	606	15,14	-	0/0/0/0	0/0/0/0
19	PGV	A	607	-	-	0/55/55/55	0/0/0/0
19	PGV	A	608	-	-	0/55/55/55	0/0/0/0
20	TGL	B	301	-	-	0/65/65/65	0/0/0/0
21	CUA	B	302	2	-	0/0/0/0	0/0/0/0
22	CHD	B	303	-	-	0/7/74/74	0/4/4/4
23	PSC	B	304	-	-	0/55/55/55	0/0/0/0
19	PGV	C	302	-	-	0/55/55/55	0/0/0/0
25	CDL	C	303	-	-	0/110/110/110	0/0/0/0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
22	CHD	C	304	-	-	0/7/74/74	0/4/4/4
22	CHD	C	305	-	-	0/7/74/74	0/4/4/4
26	PEK	C	306	-	-	0/56/56/56	0/0/0/0
19	PGV	C	307	-	-	0/55/55/55	0/0/0/0
20	TGL	D	201	-	-	0/65/65/65	0/0/0/0
26	PEK	G	101	-	-	0/56/56/56	0/0/0/0
25	CDL	G	102	-	-	1/110/110/110	0/0/0/0
26	PEK	G	103	-	-	0/56/56/56	0/0/0/0
28	DMU	J	101	-	-	0/19/59/59	0/2/2/2
22	CHD	J	102	-	-	0/7/74/74	0/4/4/4
20	TGL	L	101	-	-	0/65/65/65	0/0/0/0
28	DMU	M	101	-	-	0/19/59/59	0/2/2/2
14	HEA	N	601	1	2/2/7/16	0/24/76/76	0/0/8/8
14	HEA	N	602	1,18	3/3/7/16	0/24/76/76	0/0/8/8
18	PER	N	606	15,14	-	0/0/0/0	0/0/0/0
19	PGV	N	607	-	-	0/55/55/55	0/0/0/0
20	TGL	N	608	-	-	0/65/65/65	0/0/0/0
21	CUA	O	301	2	-	0/0/0/0	0/0/0/0
22	CHD	O	302	-	-	0/7/74/74	0/4/4/4
23	PSC	O	303	-	-	0/55/55/55	0/0/0/0
19	PGV	P	301	-	-	1/55/55/55	0/0/0/0
19	PGV	P	303	-	-	0/55/55/55	0/0/0/0
25	CDL	P	304	-	-	0/110/110/110	0/0/0/0
22	CHD	P	305	-	-	0/7/74/74	0/4/4/4
28	DMU	P	306	-	-	0/19/59/59	0/2/2/2
22	CHD	P	307	-	-	0/7/74/74	0/4/4/4
26	PEK	P	308	-	-	0/56/56/56	0/0/0/0
19	PGV	Q	201	-	-	2/55/55/55	0/0/0/0
20	TGL	Q	202	-	-	0/65/65/65	0/0/0/0
26	PEK	T	101	-	-	0/56/56/56	0/0/0/0
26	PEK	T	102	-	-	0/56/56/56	0/0/0/0
25	CDL	T	103	-	-	0/110/110/110	0/0/0/0
22	CHD	W	101	-	-	0/7/74/74	0/4/4/4
20	TGL	Y	101	-	-	0/65/65/65	0/0/0/0
28	DMU	Z	101	-	-	0/19/59/59	0/2/2/2

The worst 5 of 342 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	A	601	HEA	C18-C19	-8.00	1.12	1.33
20	B	301	TGL	OC1-CC1	-6.14	1.04	1.22
14	A	601	HEA	C16-C17	-6.08	1.32	1.53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	N	601	HEA	C18-C19	-6.01	1.17	1.33
22	O	302	CHD	C10-C5	-5.33	1.46	1.55

The worst 5 of 720 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
20	D	201	TGL	OG2-CB1-CB2	-16.27	77.78	111.55
22	W	101	CHD	C18-C13-C12	-13.80	95.04	109.08
20	Q	202	TGL	OG2-CB1-CB2	-13.01	84.54	111.55
22	P	305	CHD	C18-C13-C12	-10.06	98.84	109.08
25	C	303	CDL	C52-C51-CB5	-9.85	77.62	113.58

5 of 11 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
14	A	602	HEA	ND
14	A	602	HEA	NA
14	A	602	HEA	NB
14	N	601	HEA	ND
14	N	601	HEA	NB

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
25	G	102	CDL	CA4-OA6-CA5-C11
19	P	301	PGV	P-O11-C03-C02
19	Q	201	PGV	C02-O01-C1-O02
19	Q	201	PGV	C02-O01-C1-C2

There are no ring outliers.

40 monomers are involved in 372 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	A	601	HEA	9	0
14	A	602	HEA	1	0
18	A	606	PER	1	0
19	A	607	PGV	4	0
19	A	608	PGV	7	0
20	B	301	TGL	3	0
23	B	304	PSC	13	0
19	C	302	PGV	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
25	C	303	CDL	26	0
22	C	304	CHD	2	0
22	C	305	CHD	1	0
26	C	306	PEK	13	0
19	C	307	PGV	1	0
20	D	201	TGL	21	0
26	G	101	PEK	7	0
25	G	102	CDL	37	0
26	G	103	PEK	9	0
28	J	101	DMU	5	0
22	J	102	CHD	6	0
20	L	101	TGL	19	0
14	N	601	HEA	10	0
14	N	602	HEA	1	0
18	N	606	PER	1	0
19	N	607	PGV	1	0
20	N	608	TGL	6	0
23	O	303	PSC	19	0
19	P	301	PGV	3	0
19	P	303	PGV	6	0
25	P	304	CDL	24	0
22	P	305	CHD	6	0
28	P	306	DMU	10	0
22	P	307	CHD	1	0
26	P	308	PEK	16	0
19	Q	201	PGV	11	0
20	Q	202	TGL	15	0
26	T	101	PEK	8	0
26	T	102	PEK	10	0
25	T	103	CDL	28	0
22	W	101	CHD	6	0
20	Y	101	TGL	23	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	513/514 (99%)	-0.46	0 100 100	17, 21, 29, 68	0
1	N	513/514 (99%)	-0.38	0 100 100	18, 25, 33, 63	0
2	B	226/227 (99%)	-0.51	4 (1%) 69 74	20, 28, 49, 71	0
2	O	226/227 (99%)	-0.50	4 (1%) 69 74	25, 34, 60, 90	0
3	C	259/261 (99%)	-0.78	1 (0%) 92 94	19, 25, 37, 77	0
3	P	259/261 (99%)	-0.69	3 (1%) 79 82	20, 26, 39, 70	0
4	D	144/147 (97%)	-0.79	1 (0%) 87 90	23, 29, 44, 76	0
4	Q	144/147 (97%)	0.57	15 (10%) 7 8	29, 43, 76, 153	0
5	E	105/109 (96%)	-0.63	2 (1%) 67 72	23, 30, 53, 136	0
5	R	105/109 (96%)	-0.12	3 (2%) 52 58	26, 38, 61, 149	0
6	F	98/98 (100%)	0.06	8 (8%) 12 14	21, 31, 94, 141	0
6	S	98/98 (100%)	0.07	9 (9%) 10 11	22, 30, 80, 126	0
7	G	83/85 (97%)	0.59	17 (20%) 1 1	23, 32, 110, 158	0
7	T	83/85 (97%)	0.56	18 (21%) 1 1	23, 36, 112, 154	0
8	H	79/85 (92%)	-0.19	7 (8%) 10 12	25, 36, 92, 133	0
8	U	79/85 (92%)	-0.21	9 (11%) 6 6	31, 40, 103, 127	0
9	I	72/73 (98%)	0.26	7 (9%) 8 10	27, 41, 63, 82	0
9	V	72/73 (98%)	0.55	7 (9%) 8 10	27, 47, 69, 143	0
10	J	58/59 (98%)	0.33	7 (12%) 5 5	25, 35, 65, 134	0
10	W	58/59 (98%)	0.04	6 (10%) 7 8	27, 38, 69, 158	0
11	K	49/56 (87%)	-0.54	0 100 100	28, 35, 49, 58	0
11	X	49/56 (87%)	1.03	7 (14%) 3 3	36, 47, 68, 81	0
12	L	46/47 (97%)	-0.72	1 (2%) 62 68	22, 27, 53, 95	0
12	Y	46/47 (97%)	-0.60	1 (2%) 62 68	27, 34, 58, 125	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	M	43/46 (93%)	-0.24	4 (9%) 9 10	24, 28, 64, 118	0
13	Z	43/46 (93%)	0.04	5 (11%) 5 6	31, 38, 79, 145	0
All	All	3550/3614 (98%)	-0.29	146 (4%) 38 42	17, 29, 61, 158	0

The worst 5 of 146 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	S	97	ALA	37.2
4	Q	5	VAL	33.5
4	Q	6	VAL	14.8
6	F	97	ALA	13.2
6	F	98	HIS	11.7

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	FME	O	1	10/11	0.97	0.06	-	29,34,48,59	0
7	TPO	G	11	11/12	0.49	0.31	-	61,102,165,180	0
1	FME	A	1	10/11	0.97	0.10	-	31,39,64,89	0
2	FME	B	1	10/11	0.97	0.08	-	22,27,48,55	0
9	SAC	I	1	9/10	0.91	0.31	-	54,70,75,82	0
1	FME	N	1	10/11	0.98	0.09	-	33,40,72,74	0
9	SAC	V	1	9/10	0.52	0.37	-	114,126,154,162	0
7	TPO	T	11	11/12	0.62	0.24	-	75,96,179,187	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron



density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
22	CHD	W	101	29/29	0.45	0.43	13.85	59,136,161,164	0
22	CHD	J	102	29/29	0.57	0.44	10.79	57,133,165,172	0
25	CDL	P	304	100/100	0.86	0.27	10.35	32,78,140,149	0
25	CDL	C	303	100/100	0.85	0.32	6.72	29,72,127,137	0
18	PER	N	606	2/2	0.96	0.18	6.23	20,20,20,26	0
19	PGV	A	608	51/51	0.87	0.17	6.09	31,68,107,120	0
19	PGV	Q	201	51/51	0.85	0.26	5.71	42,79,161,191	0
20	TGL	N	608	63/63	0.86	0.20	5.68	47,76,120,136	0
25	CDL	T	103	100/100	0.69	0.28	5.27	50,91,168,202	0
20	TGL	D	201	63/63	0.85	0.14	4.99	30,55,90,101	0
25	CDL	G	102	100/100	0.70	0.30	4.80	51,96,162,187	0
18	PER	A	606	2/2	0.97	0.18	4.78	17,17,17,22	0
22	CHD	P	305	29/29	0.79	0.35	4.34	47,91,109,111	0
20	TGL	L	101	63/63	0.87	0.17	3.98	30,58,96,101	0
28	DMU	Z	101	33/33	0.85	0.20	3.69	41,47,61,63	0
20	TGL	Y	101	63/63	0.83	0.18	3.36	37,64,106,120	0
28	DMU	J	101	33/33	0.77	0.25	3.36	44,65,156,176	0
22	CHD	C	304	29/29	0.79	0.40	2.88	46,96,116,118	0
20	TGL	B	301	63/63	0.90	0.12	2.82	40,65,90,103	0
28	DMU	M	101	33/33	0.90	0.14	2.51	33,38,51,53	0
19	PGV	C	307	51/51	0.81	0.25	2.42	47,85,169,181	0
23	PSC	B	304	52/52	0.81	0.24	2.41	39,107,190,207	0
20	TGL	Q	202	63/63	0.79	0.17	2.35	40,62,99,110	0
28	DMU	P	306	33/33	0.80	0.20	2.12	36,105,192,198	0
23	PSC	O	303	52/52	0.84	0.23	1.82	37,75,191,211	0
26	PEK	T	102	53/53	0.73	0.28	1.57	44,95,159,176	0
26	PEK	G	103	53/53	0.83	0.25	1.34	46,93,150,158	0
19	PGV	P	301	51/51	0.81	0.21	1.27	45,85,147,177	0
19	PGV	C	302	51/51	0.98	0.07	1.20	22,29,81,98	0
26	PEK	P	308	53/53	0.77	0.21	1.20	38,79,120,144	0
26	PEK	C	306	53/53	0.77	0.17	1.17	34,78,128,158	0
16	MG	N	604	1/1	0.99	0.09	0.96	25,25,25,25	0
26	PEK	G	101	53/53	0.97	0.08	0.79	23,40,82,96	0
26	PEK	T	101	53/53	0.97	0.08	0.76	26,42,83,92	0
19	PGV	A	607	51/51	0.98	0.07	0.71	20,29,56,65	0
14	HEA	N	601	60/60	0.99	0.08	0.70	21,23,43,51	0
19	PGV	P	303	51/51	0.98	0.06	0.58	21,30,71,79	0
14	HEA	N	602	60/60	0.99	0.08	0.11	19,21,28,31	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
22	CHD	C	305	29/29	0.94	0.07	0.03	26,28,31,34	0
19	PGV	N	607	51/51	0.98	0.07	-0.14	22,32,67,68	0
22	CHD	O	302	29/29	0.95	0.07	-0.23	23,26,29,34	0
21	CUA	O	301	2/2	1.00	0.08	-0.25	27,27,27,27	0
14	HEA	A	601	60/60	0.99	0.07	-0.28	17,19,37,48	0
27	ZN	S	101	1/1	1.00	0.04	-0.39	26,26,26,26	0
22	CHD	B	303	29/29	0.95	0.06	-0.40	24,26,29,37	0
27	ZN	F	101	1/1	1.00	0.04	-0.51	26,26,26,26	0
21	CUA	B	302	2/2	1.00	0.08	-0.54	21,21,21,21	0
14	HEA	A	602	60/60	0.99	0.07	-0.68	17,18,25,30	0
22	CHD	P	307	29/29	0.94	0.06	-0.89	25,29,32,35	0
17	NA	N	605	1/1	1.00	0.03	-2.21	30,30,30,30	0
16	MG	A	604	1/1	0.99	0.04	-2.60	19,19,19,19	0
17	NA	A	605	1/1	1.00	0.02	-3.05	23,23,23,23	0
15	CU	A	603	1/1	1.00	0.07	-	19,19,19,19	0
24	UNX	P	302	1/1	0.81	0.39	-	38,38,38,38	0
24	UNX	C	301	1/1	0.89	0.23	-	39,39,39,39	0
15	CU	N	603	1/1	1.00	0.08	-	21,21,21,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.