



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 07:03 pm GMT

PDB ID : 3H1K  
Title : Chicken cytochrome BC1 complex with ZN<sup>++</sup> and an iodinated derivative of kresoxim-methyl bound  
Authors : Berry, E.A.; Zhang, Z.; Bellamy, H.D.; Huang, L.S.  
Deposited on : 2009-04-12  
Resolution : 3.48 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) : trunk28620

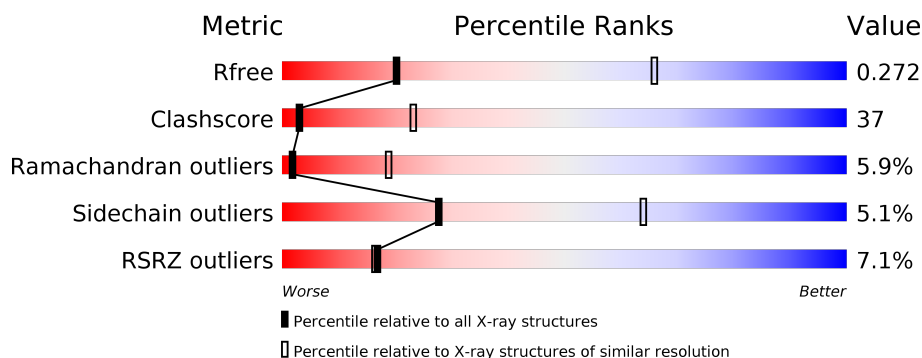
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.48 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	1049 (3.58-3.38)
Clashscore	112137	1096 (3.56-3.40)
Ramachandran outliers	110173	1063 (3.56-3.40)
Sidechain outliers	110143	1064 (3.56-3.40)
RSRZ outliers	101464	1019 (3.56-3.40)

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	446	<div> <div>2%</div> <div> <div></div> <div>44%</div> <div>50%</div> <div>5%</div> <div>..</div> </div> </div>
1	N	446	<div> <div>3%</div> <div> <div></div> <div>40%</div> <div>54%</div> <div>5%</div> <div>..</div> </div> </div>
2	B	441	<div> <div>5%</div> <div> <div></div> <div>36%</div> <div>51%</div> <div>10%</div> <div>.</div> </div> </div>
2	O	441	<div> <div>2%</div> <div> <div></div> <div>34%</div> <div>52%</div> <div>9%</div> <div>.</div> </div> </div>
3	C	380	<div> <div>2%</div> <div> <div></div> <div>44%</div> <div>50%</div> <div>5%</div> <div>.</div> </div> </div>
3	P	380	<div> <div>4%</div> <div> <div></div> <div>44%</div> <div>50%</div> <div>5%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	241	
4	Q	241	
5	E	196	
5	R	196	
6	F	110	
6	S	110	
7	G	81	
7	T	81	
8	H	77	
8	U	77	
9	I	47	
9	V	47	
10	J	61	
10	W	61	

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
11	PEE	A	2008	-	-	-	X
11	PEE	E	2005	-	-	-	X
11	PEE	P	3008	-	X	-	-
11	PEE	R	3005	-	-	-	X
12	UNL	A	3284	-	-	-	X
16	CDL	D	2003	-	-	-	X
16	CDL	Q	3003	-	-	-	X
18	GOL	C	2011	-	-	-	X
20	BOG	D	2009	-	-	-	X
20	BOG	D	2091	-	-	-	X
20	BOG	P	2010	-	-	-	X
20	BOG	Q	3009	-	-	-	X
21	FES	E	501	-	-	X	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
21	FES	R	501	-	-	X	-

## 2 Entry composition

There are 22 unique types of molecules in this entry. The entry contains 32673 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 MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	443	Total	C	N	O	S	0	0	0
			3442	2157	606	658	21			
1	N	442	Total	C	N	O	S	0	0	0
			3437	2154	605	657	21			

- Molecule 2 is a protein called MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	426	Total	C	N	O	S	0	0	0
			3164	1987	551	616	10			
2	O	422	Total	C	N	O	S	0	0	0
			3147	1977	546	614	10			

- Molecule 3 is a protein called Cytochrome b.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	380	Total	C	N	O	S	0	0	0
			3020	2024	478	505	13			
3	P	379	Total	C	N	O	S	0	0	0
			3012	2019	477	504	12			

- Molecule 4 is a protein called MITOCHONDRIAL CYTOCHROME C1, HEME PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	241	Total	C	N	O	S	0	0	0
			1898	1212	327	347	12			
4	Q	241	Total	C	N	O	S	0	0	0
			1898	1212	327	347	12			

- Molecule 5 is a protein called Cytochrome b-c1 complex subunit Rieske, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	196	Total	C	N	O	S	0	0	0
			1513	952	263	292	6			
5	R	196	Total	C	N	O	S	0	0	0
			1513	952	263	292	6			

- Molecule 6 is a protein called MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 14 KDA PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	101	Total	C	N	O	S	0	0	0
			891	570	159	159	3			
6	S	101	Total	C	N	O	S	0	0	0
			891	570	159	159	3			

- Molecule 7 is a protein called MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE UBIQUINONE-BINDING PROTEIN QP-C.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
7	G	80	Total	C	N	O	0	0	0
			672	437	119	116			
7	T	79	Total	C	N	O	0	0	0
			662	432	117	113			

- Molecule 8 is a protein called MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 11 KDA PROTEIN, COMPLEX III SUBUNIT VIII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	69	Total	C	N	O	S	0	0	0
			571	348	105	113	5			
8	U	67	Total	C	N	O	S	0	0	0
			553	338	103	107	5			

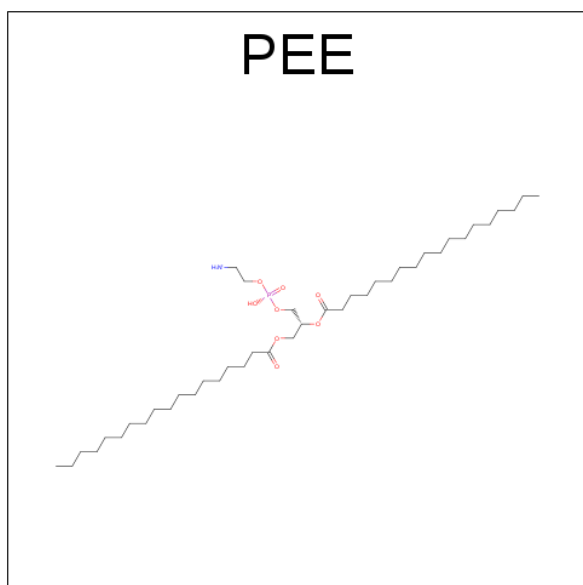
- Molecule 9 is a protein called Cytochrome b-c1 complex subunit Rieske, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	46	Total	C	N	O	S	0	0	0
			285	169	58	56	2			
9	V	44	Total	C	N	O	S	0	0	1
			275	164	56	53	2			

- Molecule 10 is a protein called MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 7.2 KDA PROTEIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
10	J	61	Total	C	N	O	0	0	0
			497	321	87	89			
10	W	60	Total	C	N	O	0	0	1
			479	311	86	82			

- Molecule 11 is PHOSPHATIDYLETHANOLAMINE (three-letter code: PEE) (formula:  $C_{41}H_{83}NO_8P$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	A	1	Total	C	O	P	0	0
			18	9	8	1		
11	C	1	Total	C	N	O	P	0
			49	39	1	8	1	
11	E	1	Total	C	N	O	P	0
			50	40	1	8	1	
11	P	1	Total	C	N	O	P	0
			49	39	1	8	1	
11	P	1	Total	O	P		0	0
			5	4	1			
11	R	1	Total	C	N	O	P	0
			50	40	1	8	1	

- Molecule 12 is UNKNOWN LIGAND (three-letter code: UNL) (formula: ).

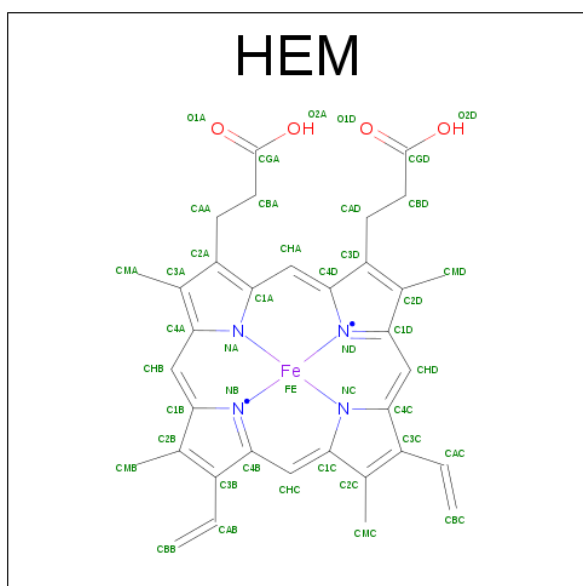
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	P	1	Total 1 O 1	0	0

*Continued on next page...*

Continued from previous page...

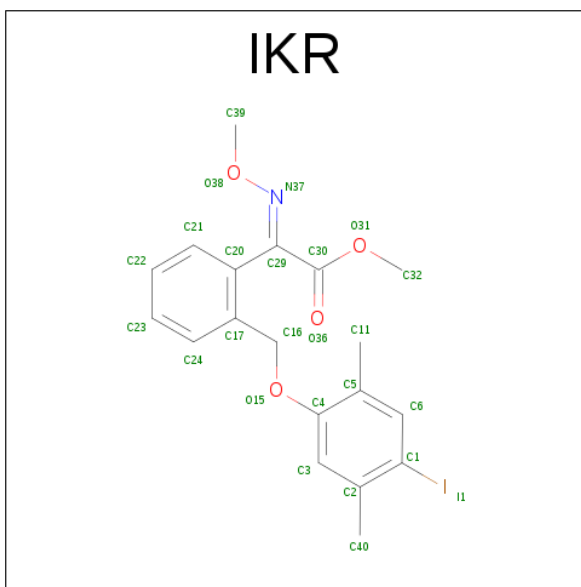
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	2	Total	O	0	0
			2	2		
12	C	1	Total	O	0	0
			1	1		
12	N	1	Total	O	0	0
			1	1		

- Molecule 13 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



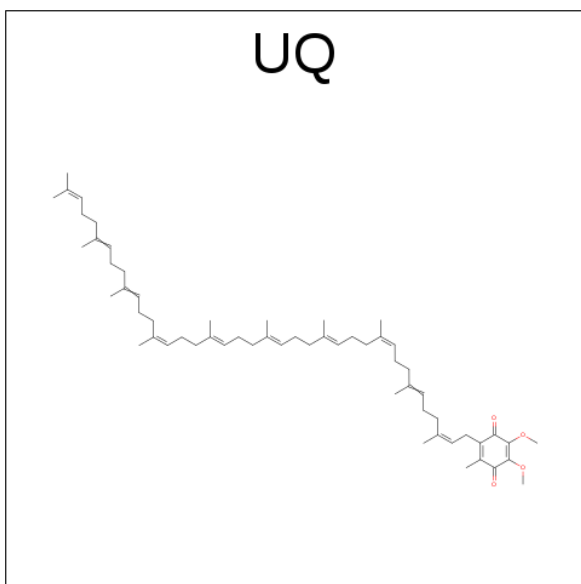
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
13	C	1	Total	C	Fe	N	O	
			43	34	1	4	4	0
13	C	1	Total	C	Fe	N	O	
			43	34	1	4	4	0
13	P	1	Total	C	Fe	N	O	
			43	34	1	4	4	0
13	P	1	Total	C	Fe	N	O	
			43	34	1	4	4	0

- Molecule 14 is METHYL (2E)-{2-[(4-iodo-2,5-dimethylphenoxy)methyl]phenyl}(methoxyimino)ethanoate (three-letter code: IKR) (formula:  $C_{19}H_{20}INO_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
14	C	1	Total	C	I	N	O	0	0
			25	19	1	1	4		
14	P	1	Total	C	I	N	O	0	0
			25	19	1	1	4		

- Molecule 15 is COENZYME Q10, (2Z,6E,10Z,14E,18E,22E,26Z)-ISOMER (three-letter code: UQ) (formula: C<sub>59</sub>H<sub>90</sub>O<sub>4</sub>).



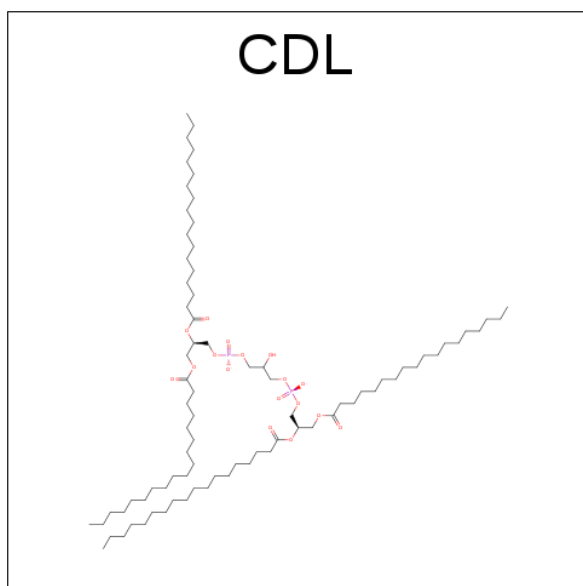
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
15	C	1	Total	C	O	0	0
			19	15	4		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
15	P	1	Total	C	O	0	0
			19	15	4		

- Molecule 16 is CARDIOLIPIN (three-letter code: CDL) (formula:  $C_{81}H_{156}O_{17}P_2$ ).

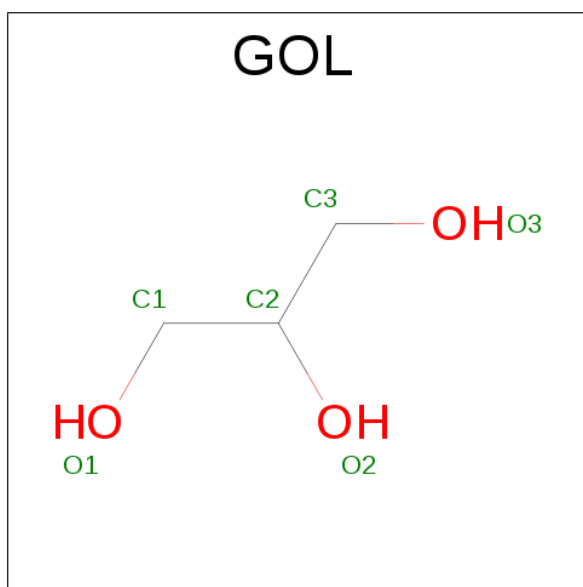


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
16	C	1	Total	C	O	P	0	0
			40	21	17	2		
16	D	1	Total	C	O	P	0	0
			42	23	17	2		
16	P	1	Total	C	O	P	0	0
			40	21	17	2		
16	Q	1	Total	C	O	P	0	0
			42	23	17	2		

- Molecule 17 is ZINC ION (three-letter code: ZN) (formula:  $Zn$ ).

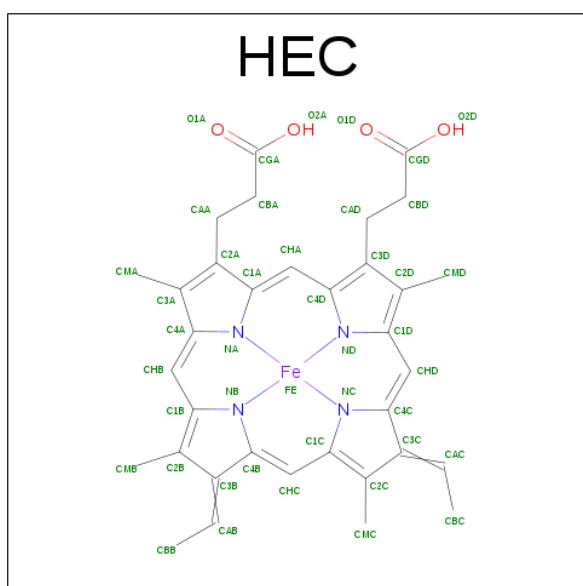
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	P	1	Total	Zn	0	0
			1	1		
17	C	1	Total	Zn	0	0
			1	1		

- Molecule 18 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



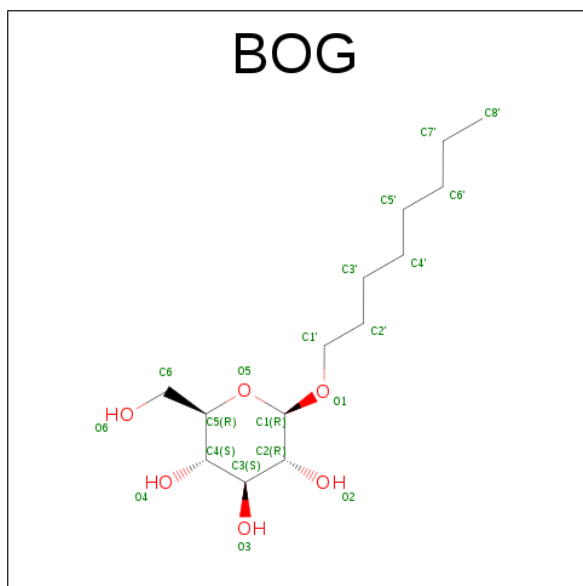
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
18	C	1	Total	C	O	0	0
			6	3	3		
18	P	1	Total	C	O	0	0
			6	3	3		

- Molecule 19 is HEME C (three-letter code: HEC) (formula:  $C_{34}H_{34}FeN_4O_4$ ).



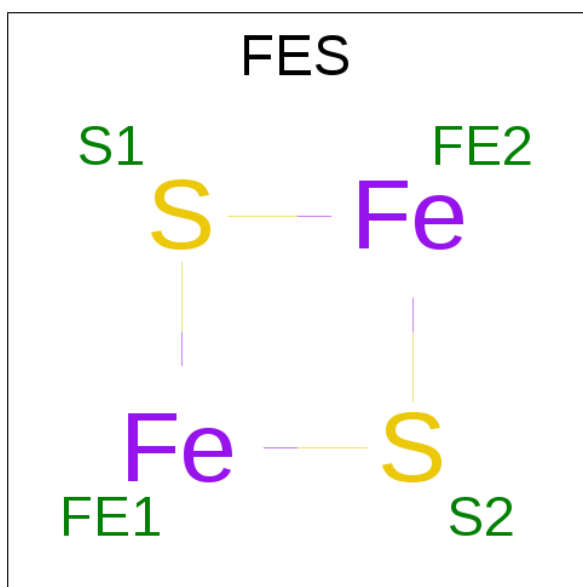
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
19	D	1	Total	C	Fe	N	O	
			43	34	1	4	4	0
19	Q	1	Total	C	Fe	N	O	
			43	34	1	4	4	0

- Molecule 20 is SUGAR (B-OCTYLGLUCOSIDE) (three-letter code: BOG) (formula:  $C_{14}H_{28}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
20	D	1	Total	C	O	0	0
			20	14	6		
20	D	1	Total	C	O	0	0
			13	7	6		
20	P	1	Total	C	O	0	0
			12	6	6		
20	P	1	Total	C	O	0	0
			13	7	6		
20	Q	1	Total	C	O	0	0
			20	14	6		

- Molecule 21 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula:  $Fe_2S_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
21	E	1	Total	Fe	S	0	0
			4	2	2		
21	R	1	Total	Fe	S	0	0
			4	2	2		

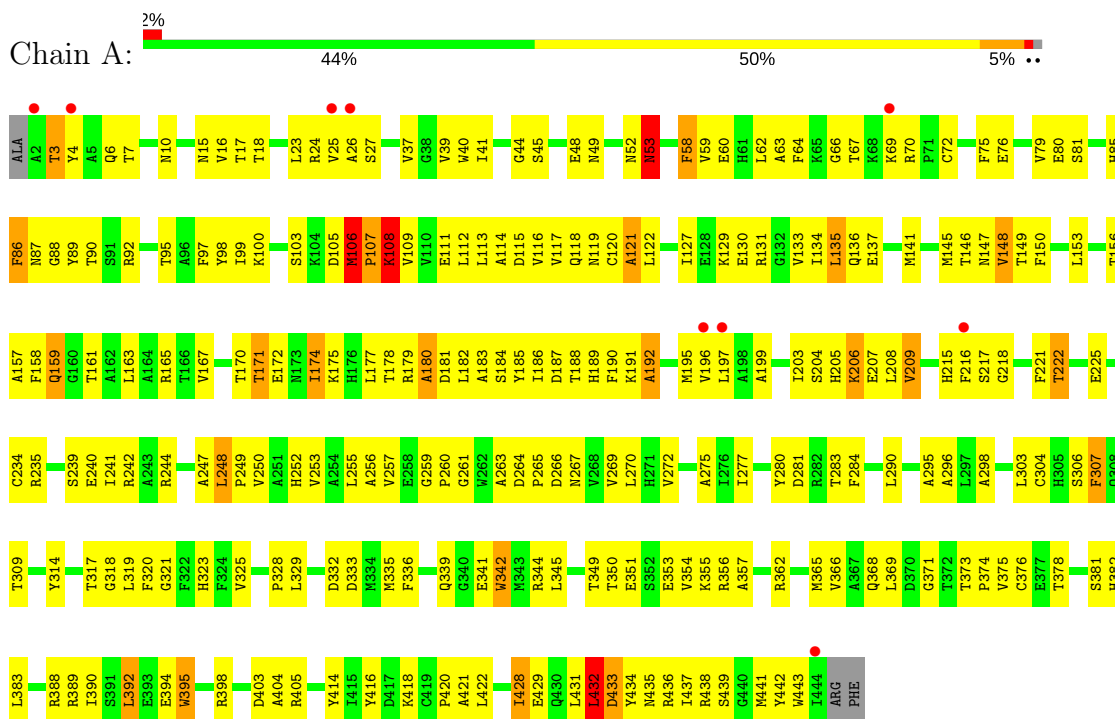
- Molecule 22 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
22	A	1	Total	O	0	0
			1	1		
22	C	7	Total	O	0	0
			7	7		
22	E	1	Total	O	0	0
			1	1		
22	P	7	Total	O	0	0
			7	7		
22	R	1	Total	O	0	0
			1	1		

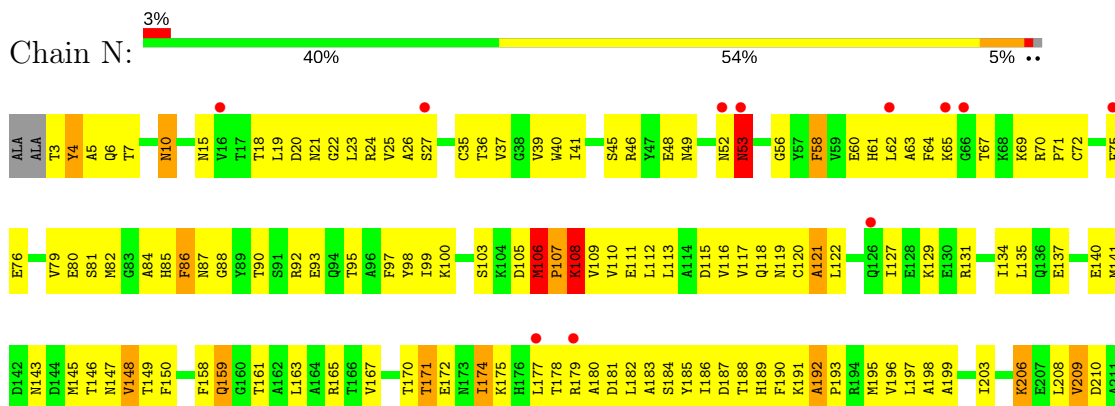
### 3 Residue-property plots

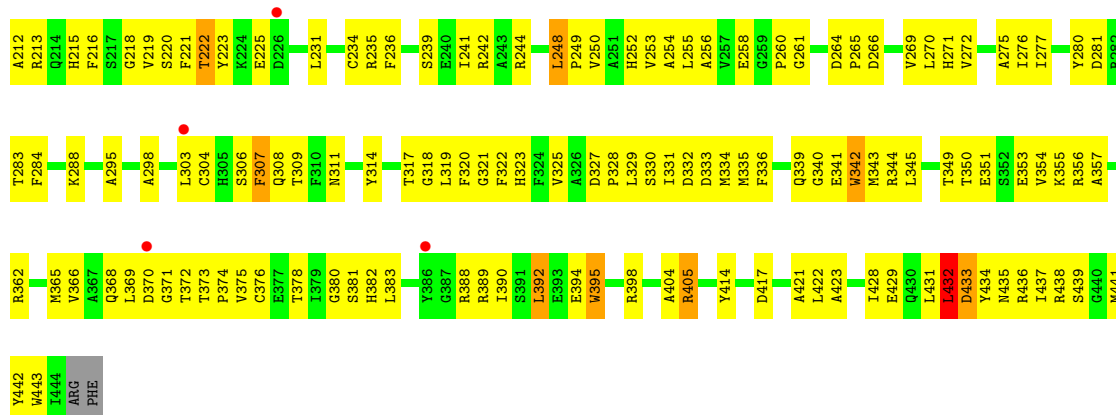
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN I

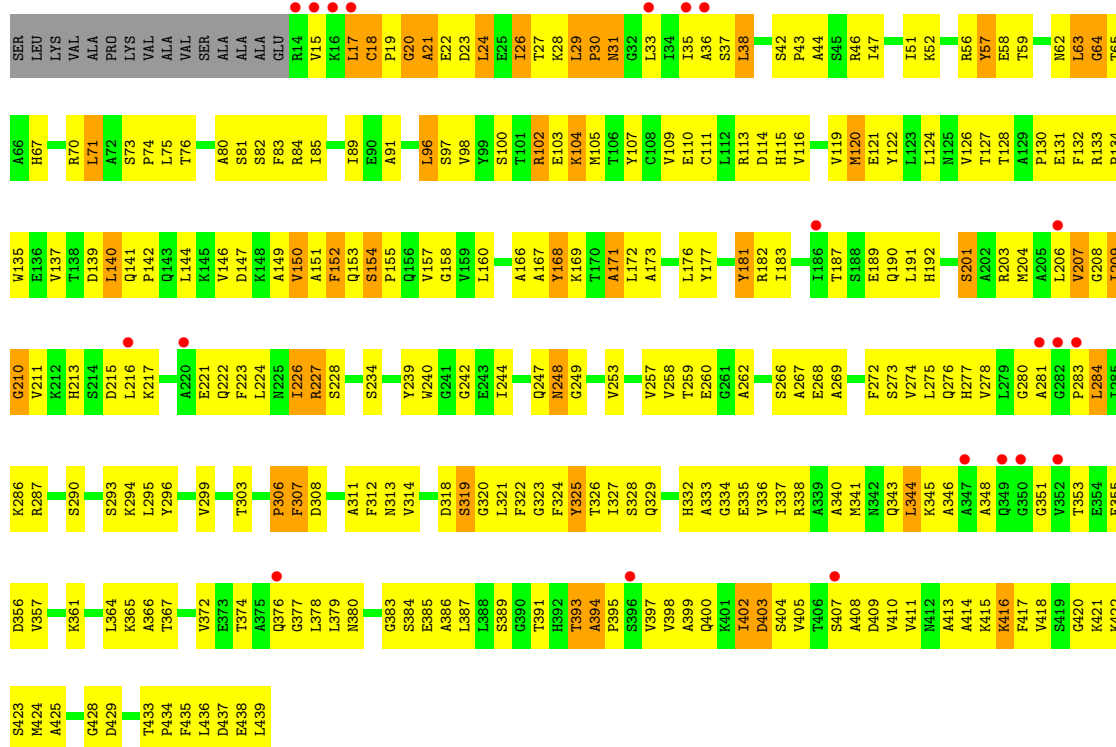


#### • Molecule 1: MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN I

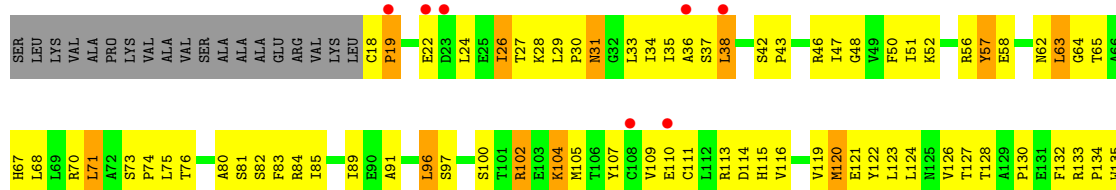


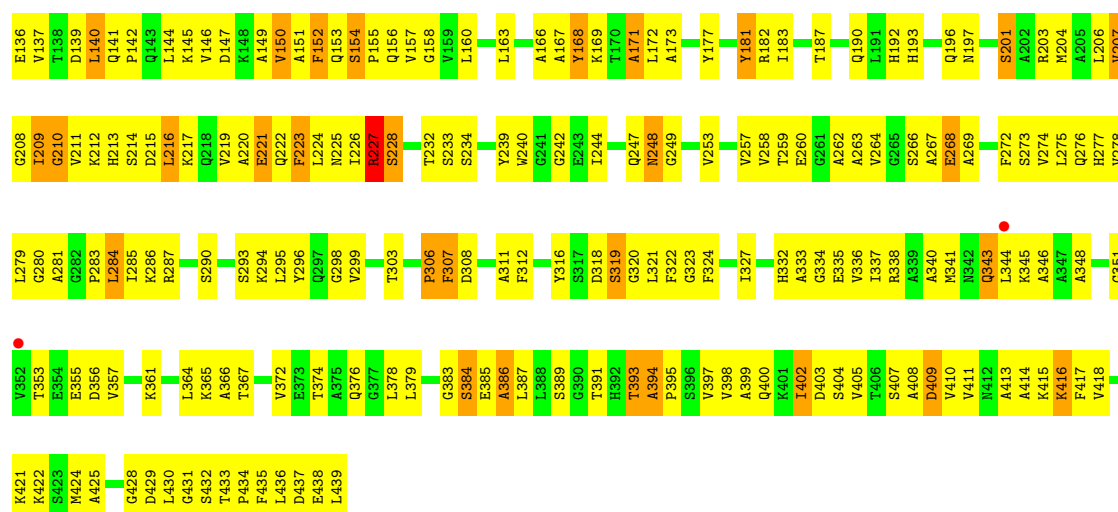


### • Molecule 2: MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN 2

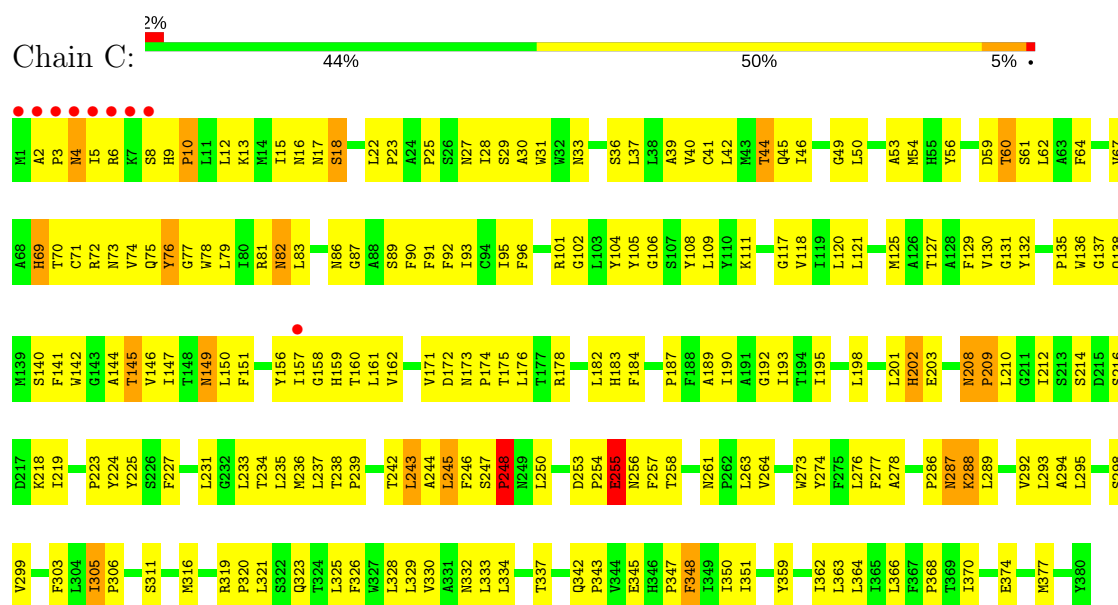


### • Molecule 2: MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN 2

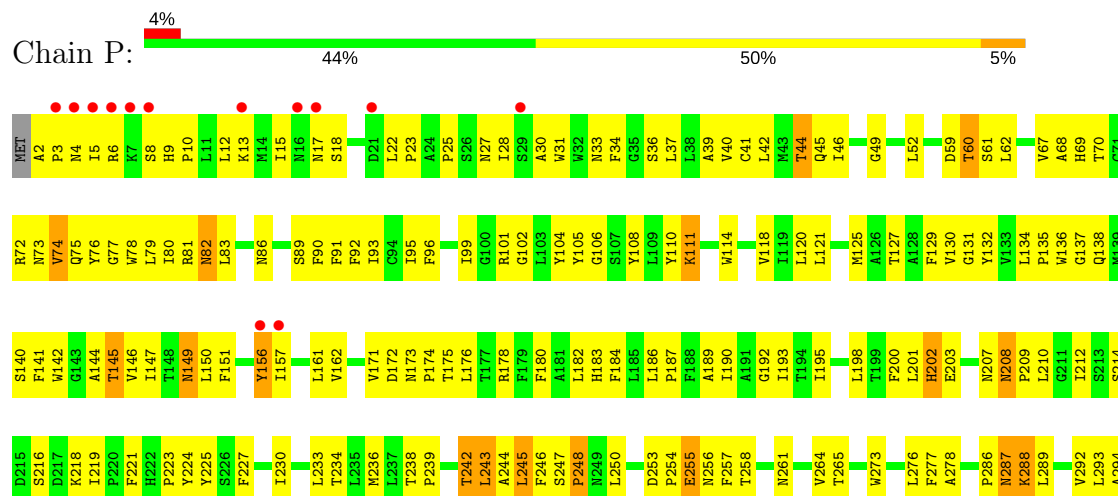


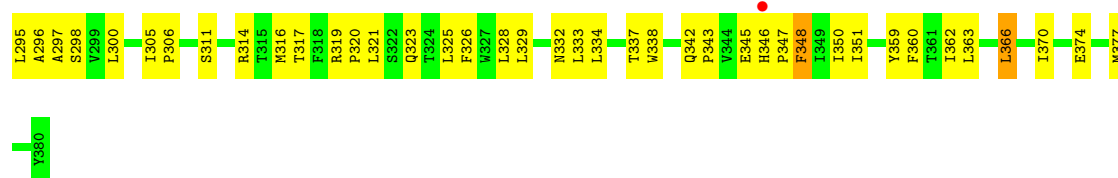


### • Molecule 3: Cytochrome b

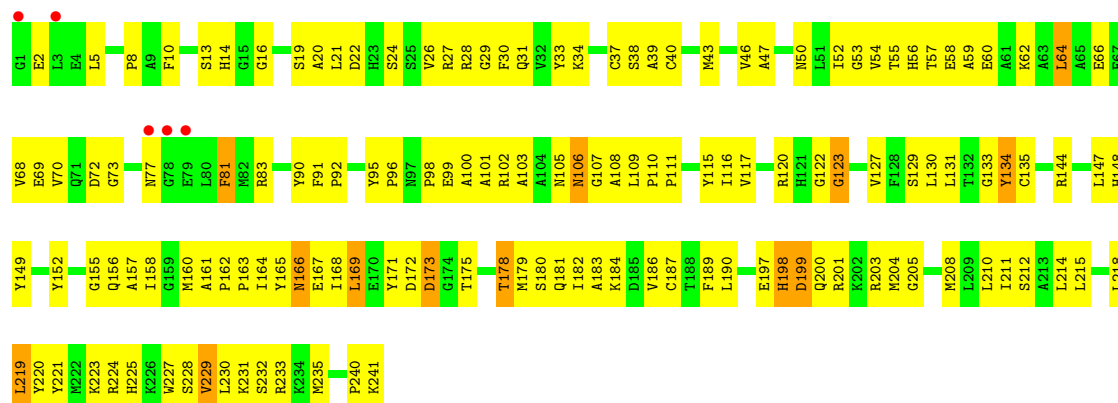


### • Molecule 3: Cytochrome b

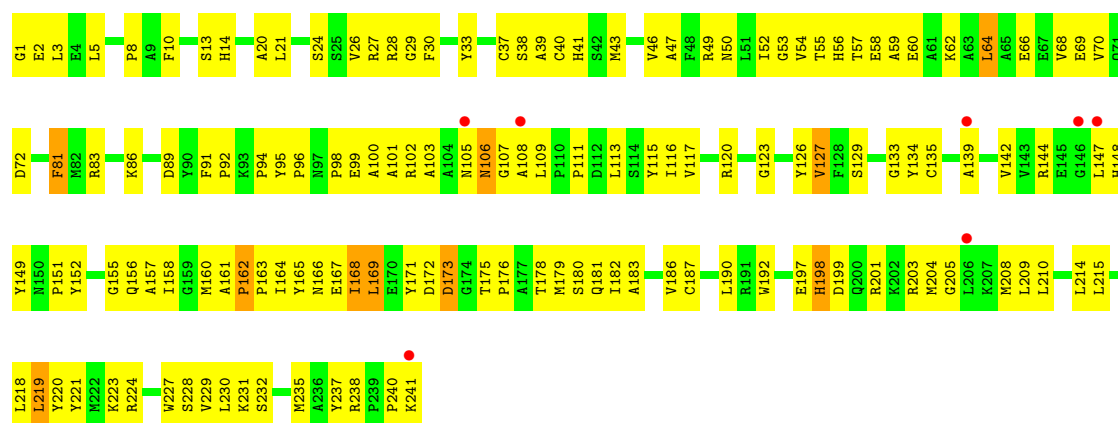
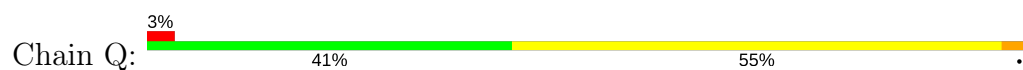




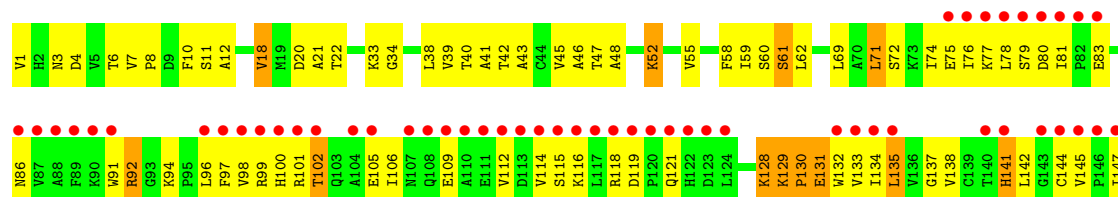
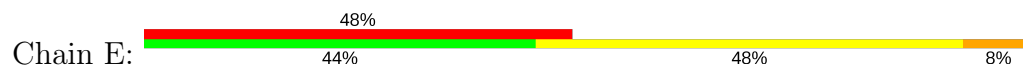
• Molecule 4: MITOCHONDRIAL CYTOCHROME C1, HEME PROTEIN

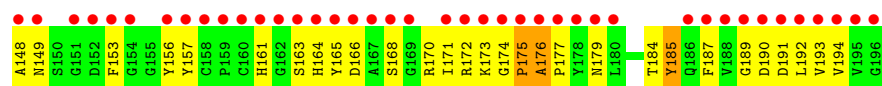


• Molecule 4: MITOCHONDRIAL CYTOCHROME C1, HEME PROTEIN

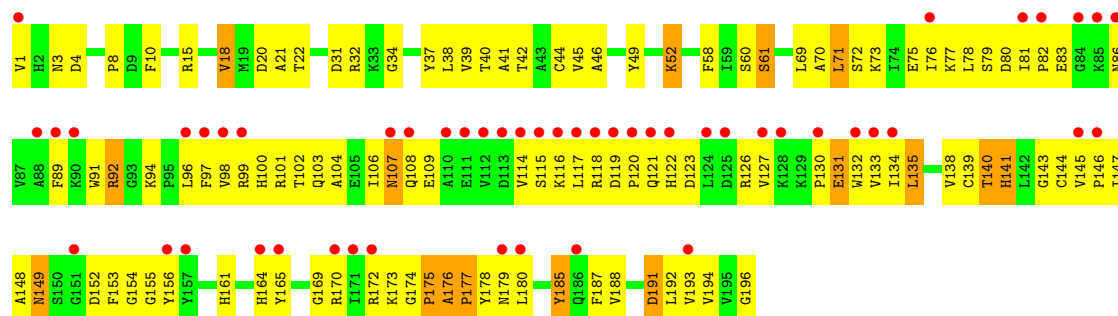
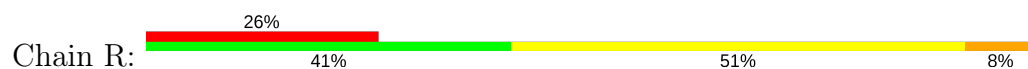


• Molecule 5: Cytochrome b-c1 complex subunit Rieske, mitochondrial





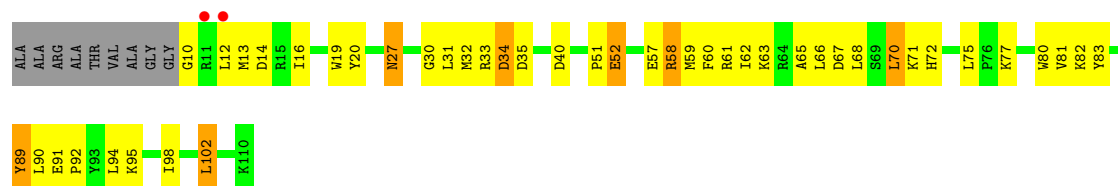
- Molecule 5: Cytochrome b-c1 complex subunit Rieske, mitochondrial



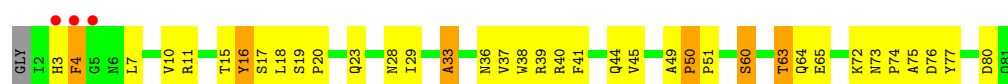
- Molecule 6: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 14 KDA PROTEIN



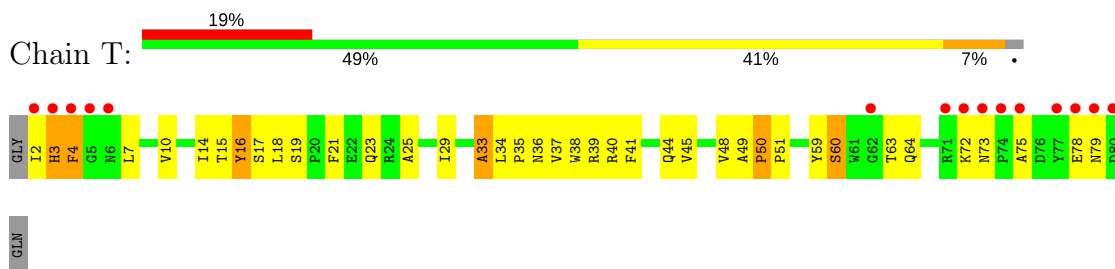
- Molecule 6: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 14 KDA PROTEIN



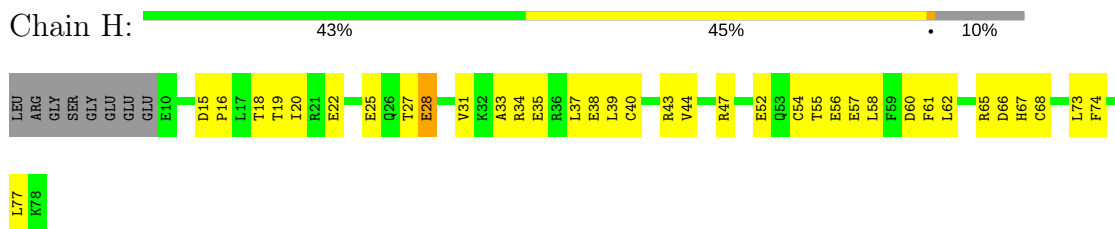
- Molecule 7: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE UBIQUINONE-BINDING PROTEIN QP-C



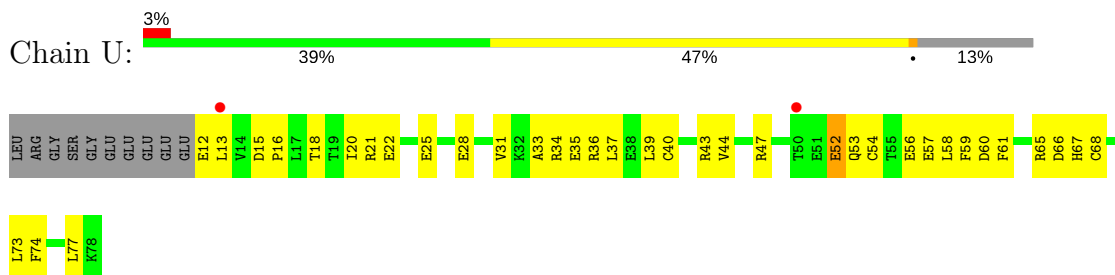
- Molecule 7: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE UBIQUINONE-BINDING PROTEIN QP-C



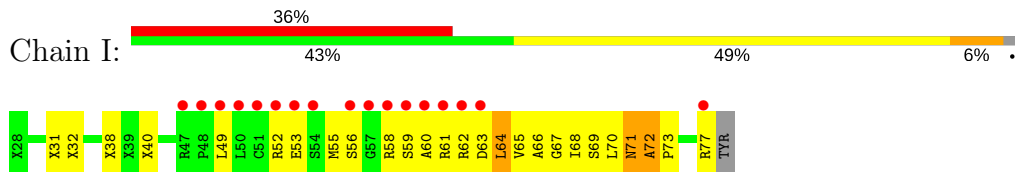
- Molecule 8: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 11 KDA PROTEIN, COMPLEX III SUBUNIT VIII



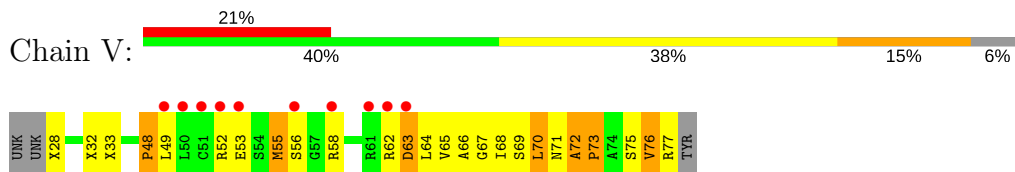
- Molecule 8: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 11 KDA PROTEIN, COMPLEX III SUBUNIT VIII



- Molecule 9: Cytochrome b-c1 complex subunit Rieske, mitochondrial



- Molecule 9: Cytochrome b-c1 complex subunit Rieske, mitochondrial



- Molecule 10: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 7.2 KDA PROTEIN





- Molecule 10: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 7.2 KDA PROTEIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	171.72Å 181.30Å 241.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	18.00 – 3.48 42.60 – 3.48	Depositor EDS
% Data completeness (in resolution range)	89.7 (18.00-3.48) 89.7 (42.60-3.48)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.20	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.72 (at 3.48Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.239 , 0.284 0.231 , 0.272	Depositor DCC
$R_{free}$ test set	2570 reflections (3.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	101.9	Xtriage
Anisotropy	0.602	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 83.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	32673	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	111.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.46% 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: GOL, ZN, IKR, CDL, UQ, FES, HEC, PEE, UNL, HEM, BOG

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	0.48	0/3513	0.70	1/4760 (0.0%)
1	N	0.50	0/3508	0.70	1/4753 (0.0%)
2	B	0.41	0/3219	0.64	0/4364
2	O	0.46	0/3202	0.67	0/4343
3	C	0.54	0/3122	0.70	0/4273
3	P	0.48	0/3114	0.67	0/4263
4	D	0.52	0/1956	0.68	0/2658
4	Q	0.44	0/1956	0.65	0/2658
5	E	0.37	0/1547	0.56	0/2103
5	R	0.39	0/1547	0.59	0/2103
6	F	0.52	0/911	0.67	0/1219
6	S	0.45	0/911	0.61	0/1219
7	G	0.54	0/694	0.72	0/941
7	T	0.49	0/684	0.71	0/929
8	H	0.50	0/579	0.68	0/775
8	U	0.40	0/561	0.60	0/751
9	I	0.37	0/218	0.64	0/293
9	V	0.40	0/218	0.66	0/293
10	J	0.50	0/508	0.60	0/682
10	W	0.43	0/490	0.61	0/660
All	All	0.47	0/32458	0.66	2/44040 (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
3	C	0	2
4	D	0	1
6	F	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	4

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	432	LEU	N-CA-C	5.68	126.33	111.00
1	N	432	LEU	N-CA-C	5.29	125.30	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	56	TYR	Sidechain
3	C	76	TYR	Sidechain
4	D	134	TYR	Sidechain
6	F	20	TYR	Sidechain

## 5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3442	0	3354	261	0
1	N	3437	0	3349	280	0
2	B	3164	0	3158	303	0
2	O	3147	0	3146	336	0
3	C	3020	0	3070	240	0
3	P	3012	0	3058	231	0
4	D	1898	0	1846	167	0
4	Q	1898	0	1846	169	0
5	E	1513	0	1478	111	0
5	R	1513	0	1478	118	0
6	F	891	0	893	50	0
6	S	891	0	893	50	0
7	G	672	0	653	38	0
7	T	662	0	645	42	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	571	0	547	33	0
8	U	553	0	535	41	0
9	I	285	0	238	37	0
9	V	275	0	238	39	0
10	J	497	0	490	28	0
10	W	479	0	478	32	0
11	A	18	0	11	0	0
11	C	49	0	72	3	0
11	E	50	0	77	0	0
11	P	54	0	72	4	0
11	R	50	0	77	0	0
12	A	2	0	0	0	0
12	C	1	0	0	0	0
12	N	1	0	0	0	0
12	P	1	0	0	0	0
13	C	86	0	60	19	0
13	P	86	0	60	20	0
14	C	25	0	20	5	0
14	P	25	0	20	4	0
15	C	19	0	17	6	0
15	P	19	0	17	6	0
16	C	40	0	24	2	0
16	D	42	0	28	1	0
16	P	40	0	24	2	0
16	Q	42	0	28	3	0
17	C	1	0	0	0	0
17	P	1	0	0	0	0
18	C	6	0	8	1	0
18	P	6	0	8	1	0
19	D	43	0	30	2	0
19	Q	43	0	30	2	0
20	D	33	0	39	1	0
20	P	25	0	22	0	0
20	Q	20	0	28	0	0
21	E	4	0	0	3	0
21	R	4	0	0	2	0
22	A	1	0	0	0	0
22	C	7	0	0	1	0
22	E	1	0	0	0	0
22	P	7	0	0	3	0
22	R	1	0	0	1	0
All	All	32673	0	32165	2414	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 37.

The worst 5 of 2414 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)
2:B:209:ILE:HG22	2:B:210:GLY:H	1.08	1.09
3:C:127:THR:HG21	13:C:501:HEM:HBB2	1.28	1.07
2:O:209:ILE:HG22	2:O:210:GLY:H	1.09	1.07
5:E:121:GLN:HG2	5:E:170:ARG:HD3	1.38	1.06
2:O:102:ARG:HG2	2:O:102:ARG:HH11	1.26	0.99

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all 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	441/446 (99%)	335 (76%)	84 (19%)	22 (5%)	2	23
1	N	440/446 (99%)	330 (75%)	87 (20%)	23 (5%)	2	21
2	B	424/441 (96%)	315 (74%)	78 (18%)	31 (7%)	1	13
2	O	420/441 (95%)	313 (74%)	78 (19%)	29 (7%)	1	15
3	C	378/380 (100%)	298 (79%)	61 (16%)	19 (5%)	2	23
3	P	377/380 (99%)	294 (78%)	67 (18%)	16 (4%)	3	28
4	D	239/241 (99%)	177 (74%)	51 (21%)	11 (5%)	3	25
4	Q	239/241 (99%)	181 (76%)	47 (20%)	11 (5%)	3	25
5	E	194/196 (99%)	119 (61%)	60 (31%)	15 (8%)	1	12
5	R	194/196 (99%)	122 (63%)	50 (26%)	22 (11%)	0	6
6	F	99/110 (90%)	74 (75%)	21 (21%)	4 (4%)	3	29
6	S	99/110 (90%)	77 (78%)	19 (19%)	3 (3%)	5	36

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	G	78/81 (96%)	59 (76%)	16 (20%)	3 (4%)	4	30
7	T	77/81 (95%)	57 (74%)	15 (20%)	5 (6%)	1	16
8	H	67/77 (87%)	44 (66%)	22 (33%)	1 (2%)	12	51
8	U	65/77 (84%)	43 (66%)	20 (31%)	2 (3%)	5	35
9	I	29/47 (62%)	15 (52%)	10 (34%)	4 (14%)	0	4
9	V	29/47 (62%)	15 (52%)	9 (31%)	5 (17%)	0	2
10	J	59/61 (97%)	38 (64%)	16 (27%)	5 (8%)	1	11
10	W	58/61 (95%)	39 (67%)	14 (24%)	5 (9%)	1	10
All	All	4006/4160 (96%)	2945 (74%)	825 (21%)	236 (6%)	2	19

5 of 236 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	159	GLN
1	A	222	THR
1	A	432	LEU
2	B	15	VAL
2	B	17	LEU

### 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	365/368 (99%)	346 (95%)	19 (5%)	27	63
1	N	365/368 (99%)	346 (95%)	19 (5%)	27	63
2	B	333/347 (96%)	308 (92%)	25 (8%)	16	51
2	O	333/347 (96%)	312 (94%)	21 (6%)	21	58
3	C	329/329 (100%)	317 (96%)	12 (4%)	40	73
3	P	328/329 (100%)	318 (97%)	10 (3%)	46	78
4	D	200/200 (100%)	191 (96%)	9 (4%)	32	68
4	Q	200/200 (100%)	191 (96%)	9 (4%)	32	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	E	166/166 (100%)	158 (95%)	8 (5%)	30	66
5	R	166/166 (100%)	158 (95%)	8 (5%)	30	66
6	F	93/96 (97%)	87 (94%)	6 (6%)	20	57
6	S	93/96 (97%)	87 (94%)	6 (6%)	20	57
7	G	71/71 (100%)	66 (93%)	5 (7%)	18	54
7	T	70/71 (99%)	67 (96%)	3 (4%)	33	69
8	H	65/71 (92%)	64 (98%)	1 (2%)	70	88
8	U	63/71 (89%)	63 (100%)	0	100	100
9	I	23/26 (88%)	22 (96%)	1 (4%)	33	69
9	V	23/26 (88%)	19 (83%)	4 (17%)	2	12
10	J	49/49 (100%)	46 (94%)	3 (6%)	22	59
10	W	47/49 (96%)	44 (94%)	3 (6%)	20	57
All	All	3382/3446 (98%)	3210 (95%)	172 (5%)	28	64

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	F	89	TYR
1	N	108	LYS
6	S	58	ARG
7	G	16	TYR
10	J	57	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 85 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	44	GLN
1	N	305	HIS
5	R	186	GLN
7	G	73	ASN
1	N	118	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 36 ligands modelled in this entry, 5 are unknown and 2 are monoatomic - leaving 29 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
11	PEE	A	2008	-	16,16,50	1.71	5 (31%)	17,20,55	0.50	0
14	IKR	C	2001	-	26,26,26	1.43	5 (19%)	30,35,35	1.50	5 (16%)
15	UQ	C	2002	-	19,19,63	2.41	10 (52%)	23,26,79	1.32	3 (13%)
16	CDL	C	2004	-	39,39,99	1.30	4 (10%)	41,51,111	1.20	5 (12%)
11	PEE	C	2007	-	48,48,50	1.25	5 (10%)	50,53,55	0.90	4 (8%)
18	GOL	C	2011	-	5,5,5	1.30	0	5,5,5	0.53	0
13	HEM	C	501	3	28,50,50	2.10	8 (28%)	17,82,82	2.92	9 (52%)
13	HEM	C	502	3	28,50,50	2.22	8 (28%)	17,82,82	1.76	5 (29%)
16	CDL	D	2003	-	41,41,99	1.15	1 (2%)	43,53,111	1.11	3 (6%)
20	BOG	D	2009	-	20,20,20	1.10	2 (10%)	25,25,25	0.88	1 (4%)
20	BOG	D	2091	-	13,13,20	1.48	3 (23%)	18,18,25	1.16	2 (11%)
19	HEC	D	501	4	28,50,50	2.22	4 (14%)	16,82,82	2.60	6 (37%)
11	PEE	E	2005	-	49,49,50	1.45	10 (20%)	51,54,55	0.99	5 (9%)
21	FES	E	501	5	0,4,4	0.00	-	0,4,4	0.00	-
20	BOG	P	2010	-	12,12,20	1.52	3 (25%)	17,17,25	0.64	0
14	IKR	P	3001	-	26,26,26	1.40	5 (19%)	30,35,35	1.53	5 (16%)
15	UQ	P	3002	-	19,19,63	2.55	10 (52%)	23,26,79	1.32	3 (13%)
16	CDL	P	3004	-	39,39,99	1.26	4 (10%)	41,51,111	1.19	5 (12%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
11	PEE	P	3007	-	48,48,50	1.22	6 (12%)	50,53,55	0.88	4 (8%)
11	PEE	P	3008	-	4,4,50	3.47	4 (100%)	6,6,55	0.56	0
18	GOL	P	3011	-	5,5,5	1.31	0	5,5,5	0.60	0
20	BOG	P	3091	-	13,13,20	1.35	2 (15%)	18,18,25	1.10	2 (11%)
13	HEM	P	501	3	28,50,50	2.34	7 (25%)	17,82,82	2.16	7 (41%)
13	HEM	P	502	3	28,50,50	2.26	8 (28%)	17,82,82	1.87	8 (47%)
16	CDL	Q	3003	-	41,41,99	1.16	1 (2%)	43,53,111	1.12	5 (11%)
20	BOG	Q	3009	-	20,20,20	1.05	2 (10%)	25,25,25	0.98	1 (4%)
19	HEC	Q	501	4	28,50,50	2.38	5 (17%)	16,82,82	2.31	5 (31%)
11	PEE	R	3005	-	49,49,50	1.49	10 (20%)	51,54,55	0.98	5 (9%)
21	FES	R	501	5	0,4,4	0.00	-	0,4,4	0.00	-

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
11	PEE	A	2008	-	-	0/19/19/54	0/0/0/0
14	IKR	C	2001	-	-	0/18/18/18	0/2/2/2
15	UQ	C	2002	-	-	0/11/35/87	0/1/1/1
16	CDL	C	2004	-	-	0/49/49/110	0/0/0/0
11	PEE	C	2007	-	-	0/52/52/54	0/0/0/0
18	GOL	C	2011	-	-	0/4/4/4	0/0/0/0
13	HEM	C	501	3	-	0/6/54/54	0/0/8/8
13	HEM	C	502	3	-	0/6/54/54	0/0/8/8
16	CDL	D	2003	-	-	0/51/51/110	0/0/0/0
20	BOG	D	2009	-	-	0/11/31/31	0/1/1/1
20	BOG	D	2091	-	-	0/4/24/31	0/1/1/1
19	HEC	D	501	4	-	0/6/54/54	0/0/8/8
11	PEE	E	2005	-	-	0/53/53/54	0/0/0/0
21	FES	E	501	5	-	0/0/4/4	0/1/1/1
20	BOG	P	2010	-	-	0/2/22/31	0/1/1/1
14	IKR	P	3001	-	-	0/18/18/18	0/2/2/2
15	UQ	P	3002	-	-	0/11/35/87	0/1/1/1
16	CDL	P	3004	-	-	0/49/49/110	0/0/0/0
11	PEE	P	3007	-	-	0/52/52/54	0/0/0/0
11	PEE	P	3008	-	-	0/0/0/54	0/0/0/0
18	GOL	P	3011	-	-	0/4/4/4	0/0/0/0
20	BOG	P	3091	-	-	0/4/24/31	0/1/1/1
13	HEM	P	501	3	-	0/6/54/54	0/0/8/8

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
13	HEM	P	502	3	-	0/6/54/54	0/0/8/8
16	CDL	Q	3003	-	-	0/51/51/110	0/0/0/0
20	BOG	Q	3009	-	-	0/11/31/31	0/1/1/1
19	HEC	Q	501	4	-	0/6/54/54	0/0/8/8
11	PEE	R	3005	-	-	0/53/53/54	0/0/0/0
21	FES	R	501	5	-	0/0/4/4	0/1/1/1

The worst 5 of 132 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
19	Q	501	HEC	C3B-C2B	-9.24	1.31	1.40
19	D	501	HEC	C3C-C2C	-6.80	1.33	1.40
19	D	501	HEC	C3B-C2B	-6.70	1.33	1.40
13	P	501	HEM	C3B-C2B	-6.38	1.31	1.40
19	Q	501	HEC	C3C-C2C	-5.57	1.34	1.40

The worst 5 of 98 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	C	501	HEM	CBA-CAA-C2A	-6.69	99.70	112.48
13	C	501	HEM	CMA-C3A-C4A	-5.77	119.59	128.46
13	P	501	HEM	CBA-CAA-C2A	-4.88	103.16	112.48
13	C	502	HEM	C4A-C3A-C2A	-4.11	104.14	107.00
13	P	502	HEM	C4A-C3A-C2A	-3.75	104.39	107.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

21 monomers are involved in 87 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	C	2001	IKR	5	0
15	C	2002	UQ	6	0
16	C	2004	CDL	2	0
11	C	2007	PEE	3	0
18	C	2011	GOL	1	0
13	C	501	HEM	6	0
13	C	502	HEM	13	0
16	D	2003	CDL	1	0
20	D	2091	BOG	1	0
19	D	501	HEC	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
21	E	501	FES	3	0
14	P	3001	IKR	4	0
15	P	3002	UQ	6	0
16	P	3004	CDL	2	0
11	P	3007	PEE	4	0
18	P	3011	GOL	1	0
13	P	501	HEM	9	0
13	P	502	HEM	11	0
16	Q	3003	CDL	3	0
19	Q	501	HEC	2	0
21	R	501	FES	2	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	443/446 (99%)	-0.09	9 (2%) 65 59	56, 103, 147, 160	0
1	N	442/446 (99%)	-0.06	15 (3%) 46 40	65, 100, 132, 147	0
2	B	426/441 (96%)	0.19	21 (4%) 30 26	85, 126, 182, 218	0
2	O	422/441 (95%)	-0.05	9 (2%) 64 57	70, 108, 147, 161	0
3	C	380/380 (100%)	-0.30	9 (2%) 59 52	40, 64, 131, 204	0
3	P	379/380 (99%)	-0.11	14 (3%) 42 37	52, 97, 148, 187	0
4	D	241/241 (100%)	-0.22	5 (2%) 64 57	48, 72, 121, 138	0
4	Q	241/241 (100%)	-0.00	7 (2%) 52 45	74, 108, 149, 167	0
5	E	196/196 (100%)	2.73	94 (47%) 0 0	57, 214, 260, 262	125 (63%)
5	R	196/196 (100%)	1.21	51 (26%) 1 1	64, 183, 229, 234	1 (0%)
6	F	101/110 (91%)	-0.35	1 (0%) 82 76	56, 75, 97, 117	0
6	S	101/110 (91%)	0.03	2 (1%) 65 59	87, 124, 174, 190	0
7	G	80/81 (98%)	0.07	3 (3%) 41 36	58, 86, 158, 166	0
7	T	79/81 (97%)	0.60	15 (18%) 1 2	81, 126, 222, 227	0
8	H	69/77 (89%)	-0.25	0 100 100	70, 98, 119, 154	0
8	U	67/77 (87%)	0.21	2 (2%) 51 44	133, 165, 209, 211	0
9	I	31/47 (65%)	3.40	17 (54%) 0 0	147, 185, 257, 259	0
9	V	31/47 (65%)	1.80	10 (32%) 0 0	121, 163, 242, 244	0
10	J	61/61 (100%)	0.09	2 (3%) 47 41	75, 93, 139, 177	0
10	W	60/61 (98%)	0.38	2 (3%) 47 41	84, 108, 157, 165	0
All	All	4046/4160 (97%)	0.19	288 (7%) 17 16	40, 104, 204, 262	126 (3%)

The worst 5 of 288 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
9	I	51	CYS	25.5
5	E	162	GLY	15.2
5	E	172	ARG	12.9
5	E	117	LEU	12.8
5	E	171	ILE	12.6

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

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

## 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( $\text{\AA}^2$ )	Q<0.9
11	PEE	A	2008	18/51	0.66	0.58	7.12	89,197,198,199	0
11	PEE	E	2005	50/51	0.78	0.52	4.55	113,127,137,139	0
18	GOL	C	2011	6/6	0.94	0.34	4.42	69,72,75,78	0
11	PEE	R	3005	50/51	0.68	0.48	4.23	114,131,139,141	0
20	BOG	D	2091	13/20	0.63	0.47	3.78	191,192,193,193	0
16	CDL	Q	3003	42/100	0.71	0.44	3.68	183,193,208,208	0
20	BOG	Q	3009	20/20	0.82	0.52	3.08	103,118,120,121	0
20	BOG	P	2010	12/20	0.41	0.55	2.97	248,249,250,250	0
20	BOG	D	2009	20/20	0.90	0.40	2.64	93,104,108,108	0
16	CDL	D	2003	42/100	0.87	0.34	2.63	134,141,153,153	0
11	PEE	P	3007	49/51	0.89	0.32	1.98	102,116,140,141	0
12	UNL	A	3284	1/-	0.83	0.49	1.90	31,31,31,31	0
15	UQ	C	2002	19/63	0.91	0.27	1.81	107,111,115,115	0
18	GOL	P	3011	6/6	0.87	0.29	1.66	103,108,109,110	0
15	UQ	P	3002	19/63	0.86	0.36	1.62	134,146,148,149	0
16	CDL	P	3004	40/100	0.76	0.38	1.55	147,157,162,162	0
16	CDL	C	2004	40/100	0.89	0.29	1.50	103,110,126,128	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
11	PEE	C	2007	49/51	0.93	0.24	1.12	68,80,102,104	0
13	HEM	P	501	43/43	0.98	0.25	0.94	65,72,79,81	0
14	IKR	P	3001	25/25	0.97	0.23	0.80	86,89,108,116	0
13	HEM	C	501	43/43	0.98	0.23	0.75	48,52,62,66	0
13	HEM	C	502	43/43	0.99	0.21	0.26	41,46,56,61	0
19	HEC	D	501	43/43	0.98	0.20	0.21	57,60,68,70	0
13	HEM	P	502	43/43	0.98	0.20	-0.11	68,73,82,82	0
14	IKR	C	2001	25/25	0.99	0.17	-0.24	65,67,75,79	0
19	HEC	Q	501	43/43	0.96	0.21	-0.57	87,91,97,99	0
21	FES	E	501	4/4	0.92	0.36	-0.89	257,257,257,257	4
21	FES	R	501	4/4	0.90	0.06	-1.38	201,201,201,202	0
12	UNL	C	4234	1/-	0.53	0.99	-	58,58,58,58	0
12	UNL	P	4236	1/-	0.69	0.76	-	94,94,94,94	0
17	ZN	C	2012	1/1	0.94	0.05	-	104,104,104,104	0
17	ZN	P	3012	1/1	0.98	0.04	-	121,121,121,121	0
11	PEE	P	3008	5/51	0.83	0.38	-	184,184,184,185	0
12	UNL	N	4231	1/-	0.37	0.59	-	90,90,90,90	0
20	BOG	P	3091	13/20	0.28	0.83	-	298,299,299,300	0
12	UNL	A	3231	1/-	0.75	0.82	-	45,45,45,45	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.