wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 10:14 AM GMT

PDB ID	:	3L72
Title	:	Chicken cytochrome BC1 complex with kresoxym-I-dimethyl bound
Authors	:	Huang, L.; Zhang, Z.; Berry, E.A.
Deposited on	:	2009-12-27
Resolution	:	3.06 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

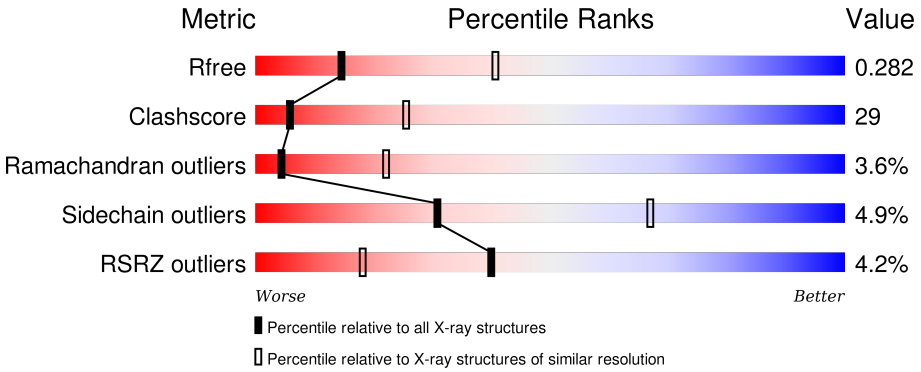
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.06 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.



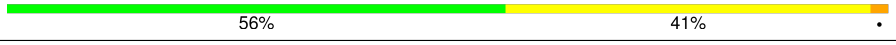
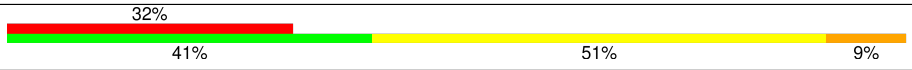
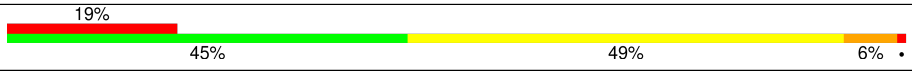
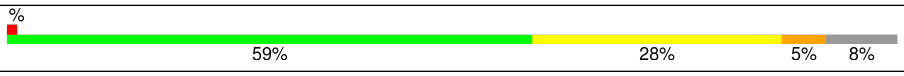
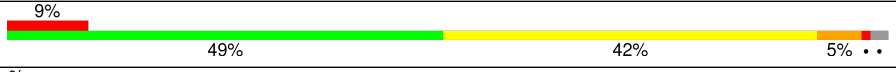
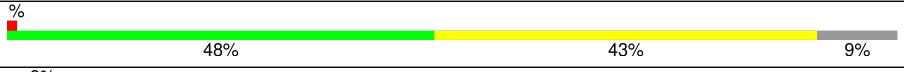
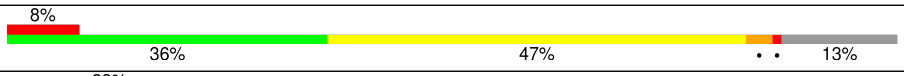
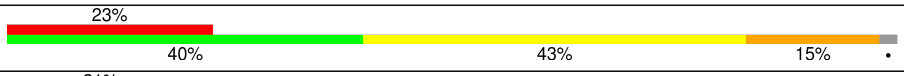
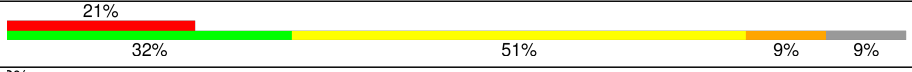
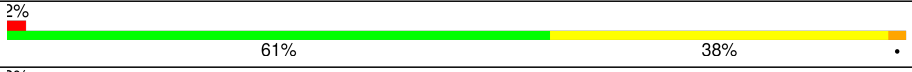
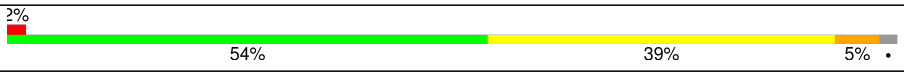
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1191 (3.10-3.02)
Clashscore	102246	1303 (3.10-3.02)
Ramachandran outliers	100387	1254 (3.10-3.02)
Sidechain outliers	100360	1254 (3.10-3.02)
RSRZ outliers	91569	1197 (3.10-3.02)

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 >=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 <=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	54% 41%
1	N	446	49% 45% 5%
2	B	441	41% 48% 6% 5%
2	O	441	44% 47% 5%
3	C	380	56% 40%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	P	380	
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
13	UQ	C	2002	-	-	-	X
13	UQ	P	3002	-	-	-	X
14	CDL	D	2003	-	-	-	X
14	CDL	P	3004	-	-	-	X
14	CDL	Q	3003	-	-	-	X
15	PEE	C	2007	-	-	-	X
15	PEE	C	2008	-	-	-	X
15	PEE	E	2005	-	-	-	X
15	PEE	P	3007	-	-	-	X
15	PEE	P	3008	-	X	-	-
15	PEE	R	3005	-	-	-	X
16	GOL	C	2011	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
16	GOL	P	3011	-	-	-	X
18	BOG	D	2009	-	-	-	X
18	BOG	P	2010	-	-	-	X
18	BOG	Q	3009	-	-	-	X
19	FES	E	501	-	-	X	-
19	FES	R	501	-	-	X	-

2 Entry composition

There are 20 unique types of molecules in this entry. The entry contains 32648 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	444	Total	C	N	O	S	0	0	0
			3447	2160	607	659	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	421	Total	C	N	O	S	0	0	0
			3137	1970	545	613	9			
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
			3017	2022	478	505	12			
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 5, RIESKE IRONSULFUR PROTEIN, 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
			1509	950	263	290	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	70	Total	C	N	O	S	0	0	0
			574	350	105	114	5			
8	U	67	Total	C	N	O	S	0	0	0
			553	338	103	107	5			

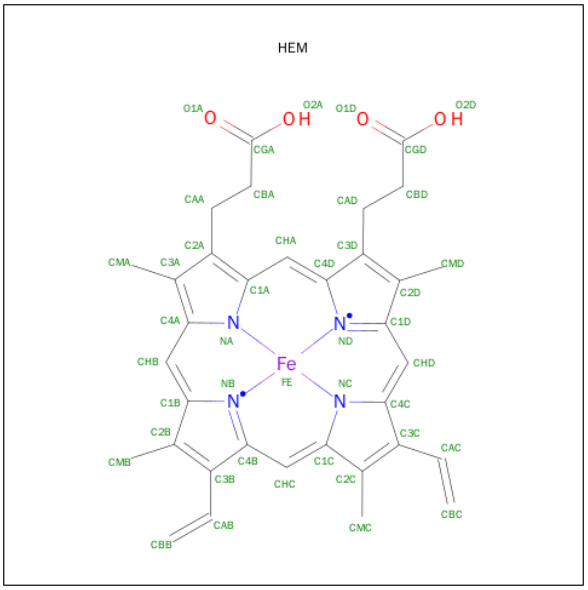
- Molecule 9 is a protein called CYTOCHROME B-C1 COMPLEX SUBUNIT RIESKE, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	46	Total	C	N	O	S	0	0	0
			287	171	58	56	2			
9	V	43	Total	C	N	O	S	0	0	0
			277	167	55	53	2			

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

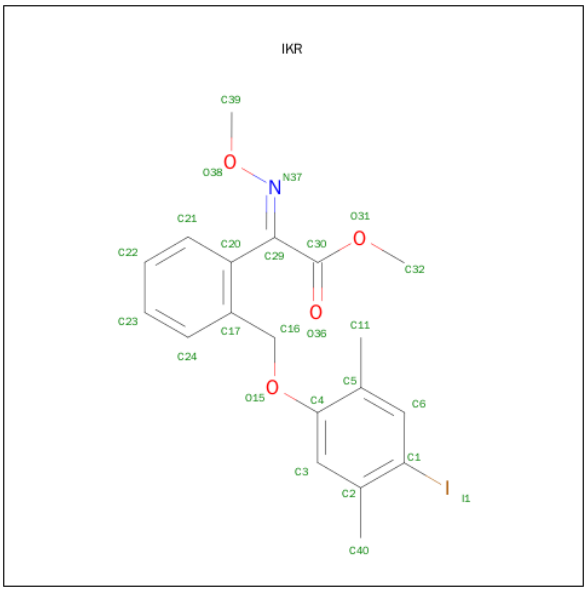
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
10	J	61	Total 497	C 321	N 87	O 89	0	0	0
10	W	60	Total 479	C 311	N 86	O 82	0	0	1

- Molecule 11 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



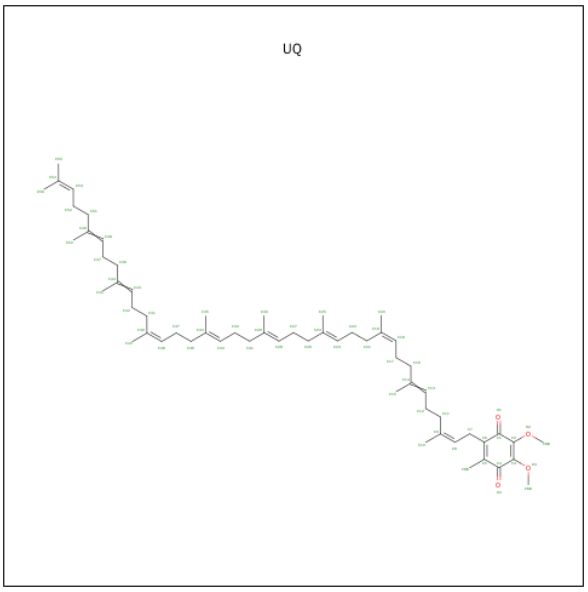
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
11	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
11	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
11	P	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
11	P	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

- Molecule 12 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
12	C	1	Total	C	I	N	O	0	0
			25	19	1	1	4		
12	P	1	Total	C	I	N	O	0	0
			25	19	1	1	4		

- Molecule 13 is COENZYME Q10, (2Z,6E,10Z,14E,18E,22E,26Z)-ISOMER (three-letter code: UQ) (formula: $C_{59}H_{90}O_4$).



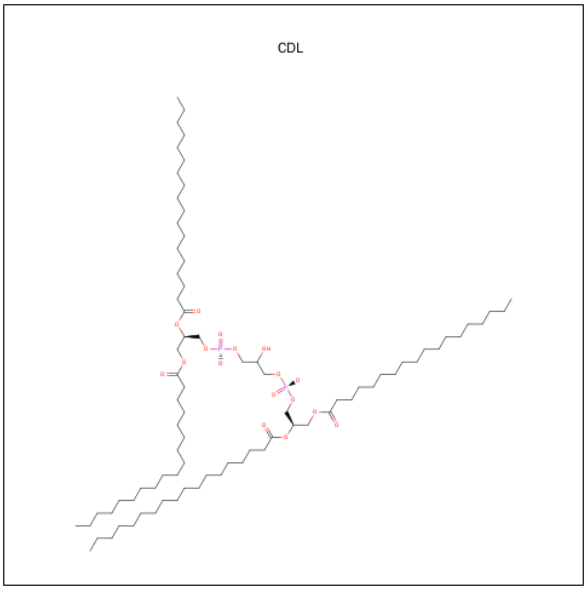
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	C	1	Total	C	O	0	0
			19	15	4		

Continued on next page...

Continued from previous page...

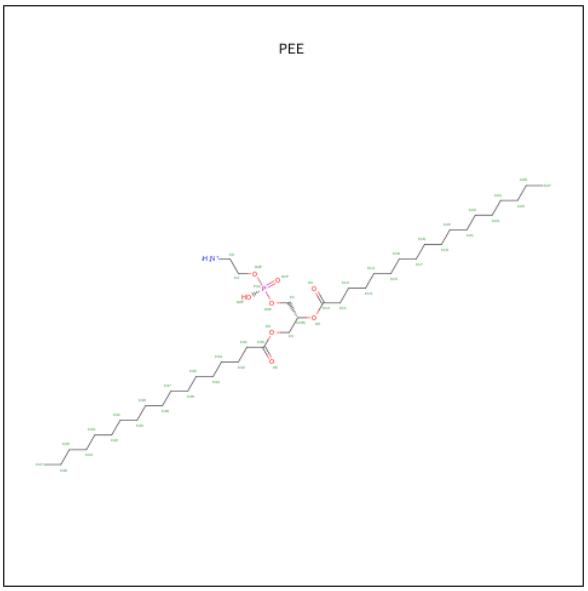
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	P	1	Total	C	O	0	0
			19	15	4		

- Molecule 14 is CARDIOLIPIN (three-letter code: CDL) (formula: $C_{81}H_{156}O_{17}P_2$).



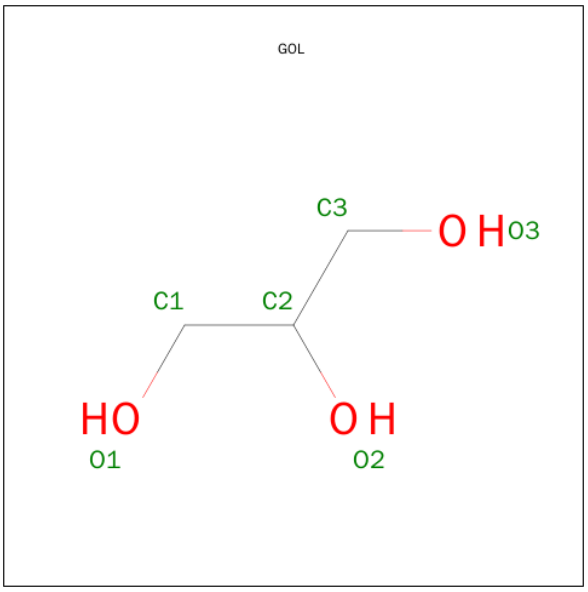
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
14	C	1	Total	C	O	P	0	0
			40	21	17	2		
14	D	1	Total	C	O	P	0	0
			42	23	17	2		
14	P	1	Total	C	O	P	0	0
			40	21	17	2		
14	Q	1	Total	C	O	P	0	0
			42	23	17	2		

- Molecule 15 is 1,2-DIOLEOYL-SN-GLYCERO-3-PHOSPHOETHANOLAMINE (three-letter code: PEE) (formula: $C_{41}H_{83}NO_8P$).



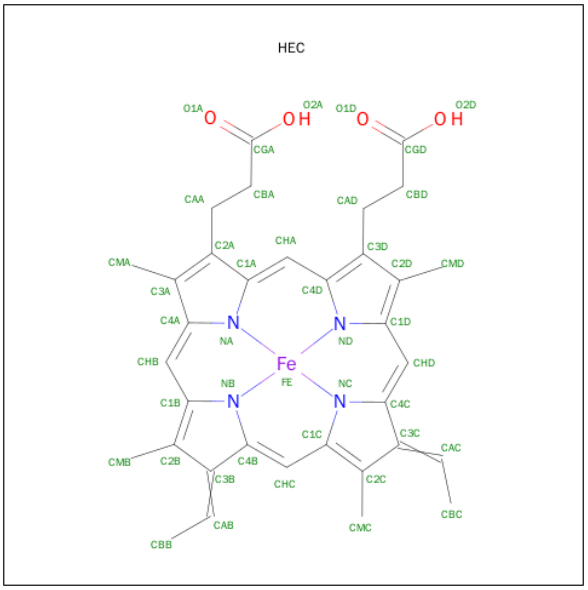
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	C	1	Total C N O P 49 39 1 8 1	0	0
15	C	1	Total C O P 21 12 8 1	0	0
15	E	1	Total C N O P 50 40 1 8 1	0	0
15	P	1	Total C N O P 49 39 1 8 1	0	0
15	P	1	Total O P 5 4 1	0	0
15	R	1	Total C N O P 50 40 1 8 1	0	0

- Molecule 16 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



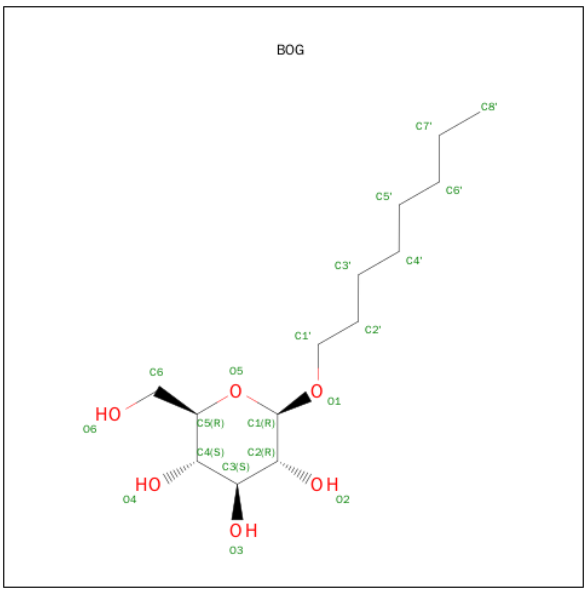
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	C	1	Total C O 6 3 3	0	0
16	P	1	Total C O 6 3 3	0	0

- Molecule 17 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).



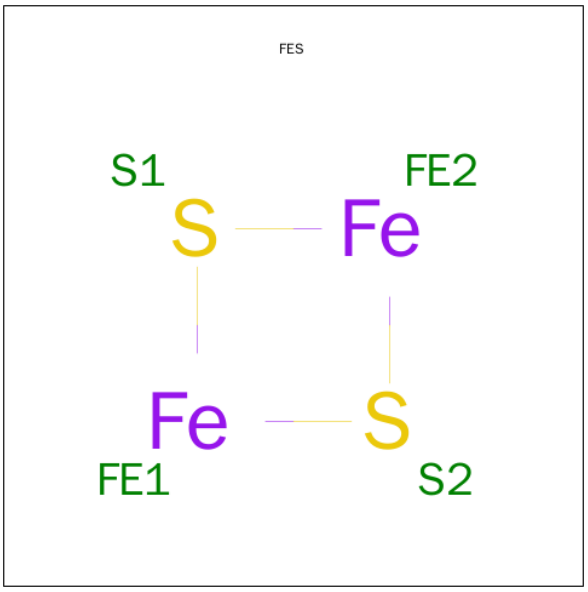
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	D	1	Total C Fe N O 43 34 1 4 4	0	0
17	Q	1	Total C Fe N O 43 34 1 4 4	0	0

- Molecule 18 is SUGAR (B-OCTYLGLUCOSIDE) (three-letter code: BOG) (formula: $C_{14}H_{28}O_6$).



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

- Molecule 19 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).



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

- Molecule 20 is water.

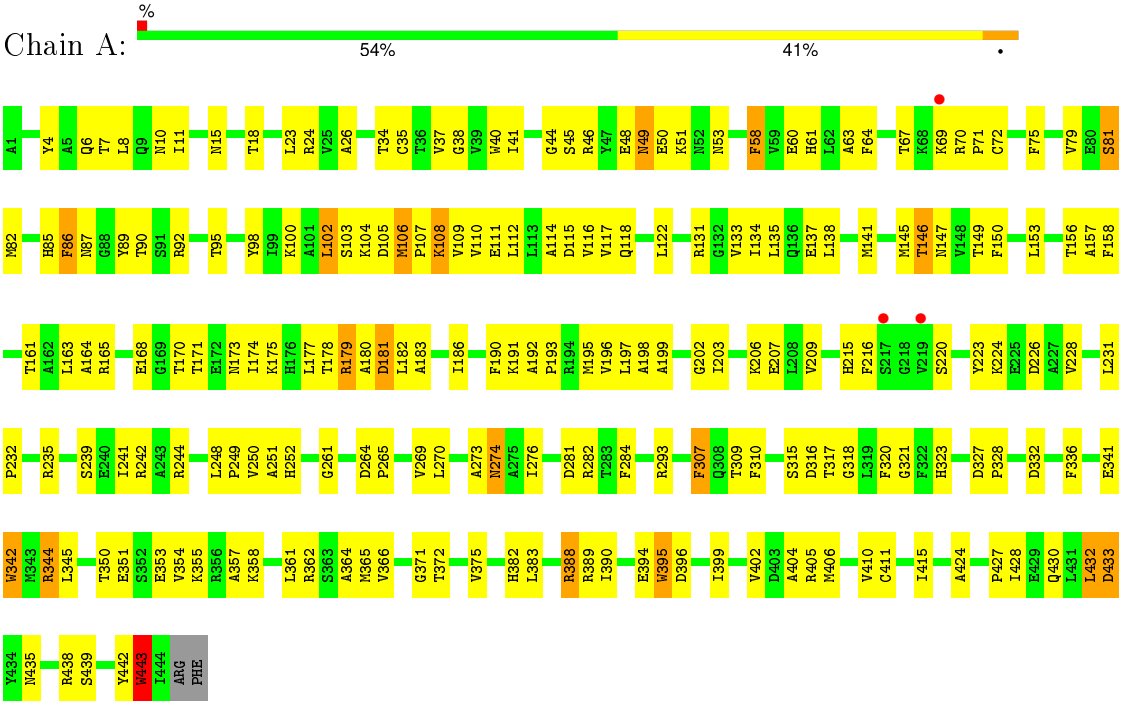
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
20	C	8	Total O 8 8	0	0
20	E	1	Total O 1 1	0	0
20	P	8	Total O 8 8	0	0
20	R	1	Total O 1 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

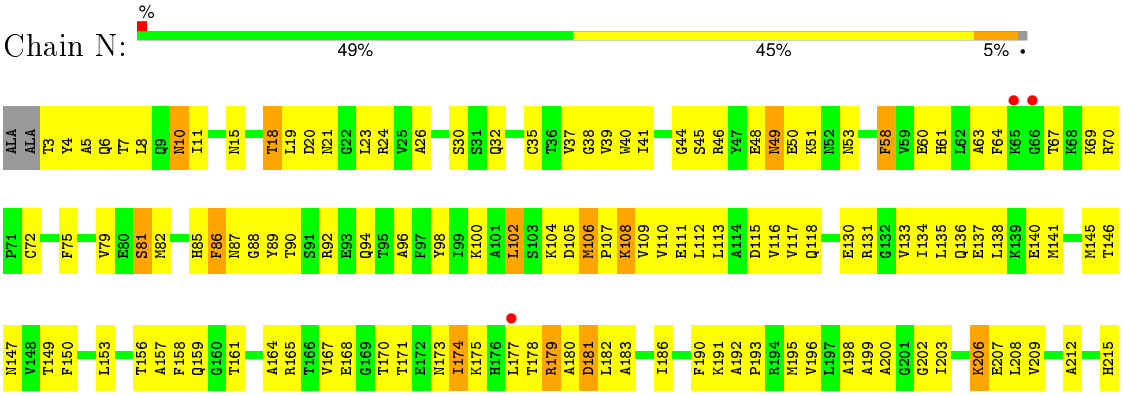
• Molecule 1: MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN I

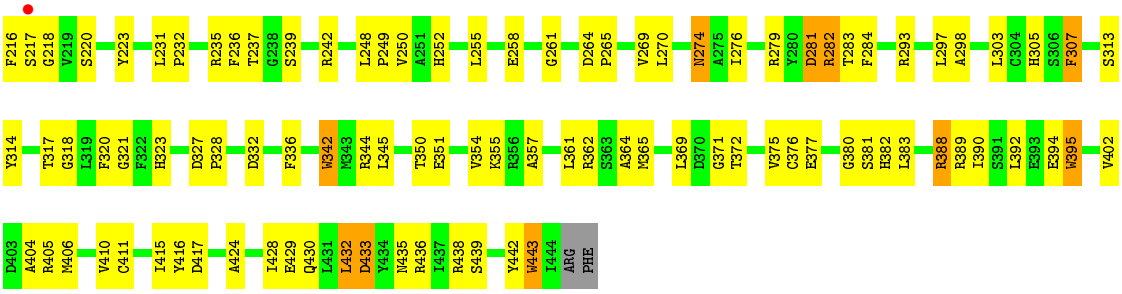
Chain A:



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

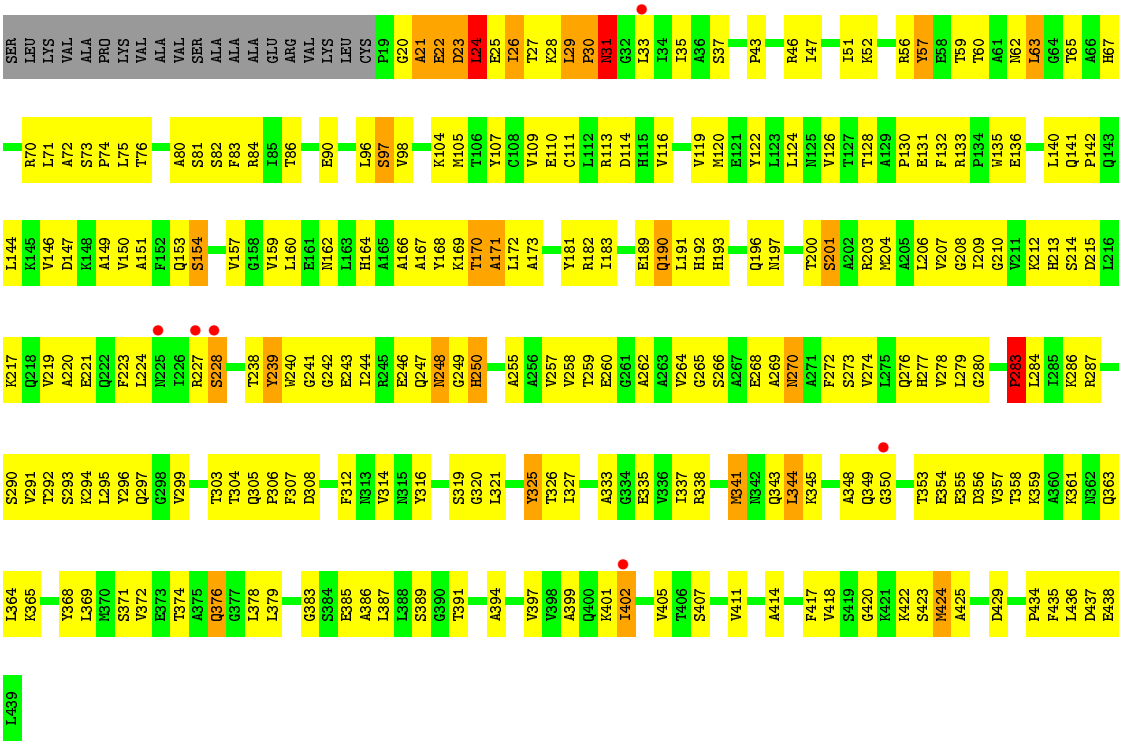
Chain N:





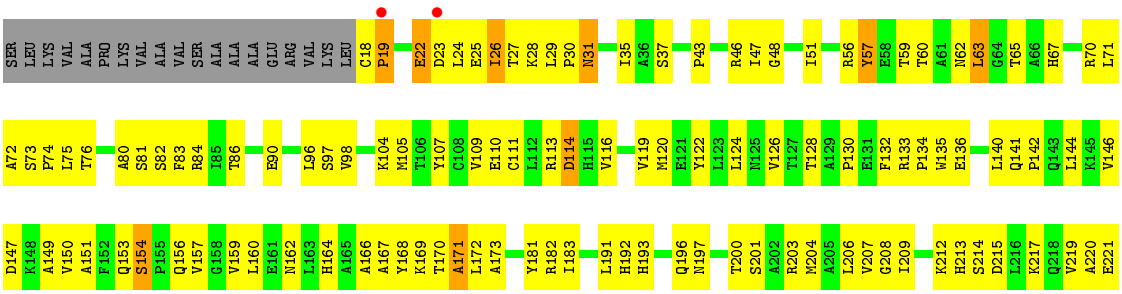
• Molecule 2: MITOCHONDRIAL UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN 2

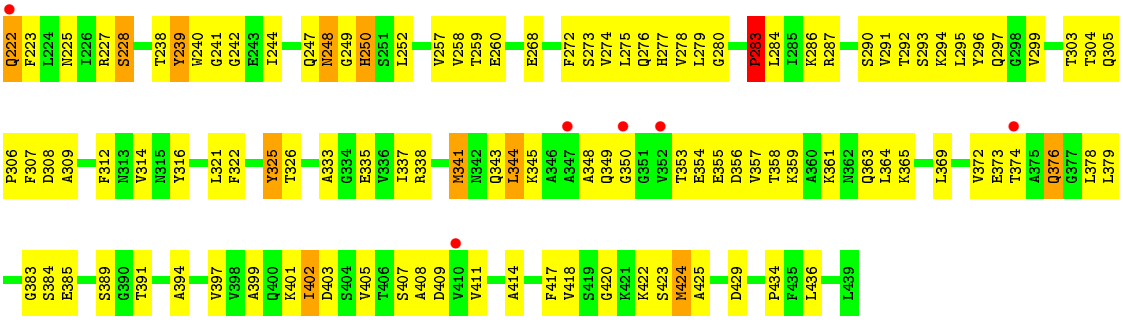
Chain B:



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

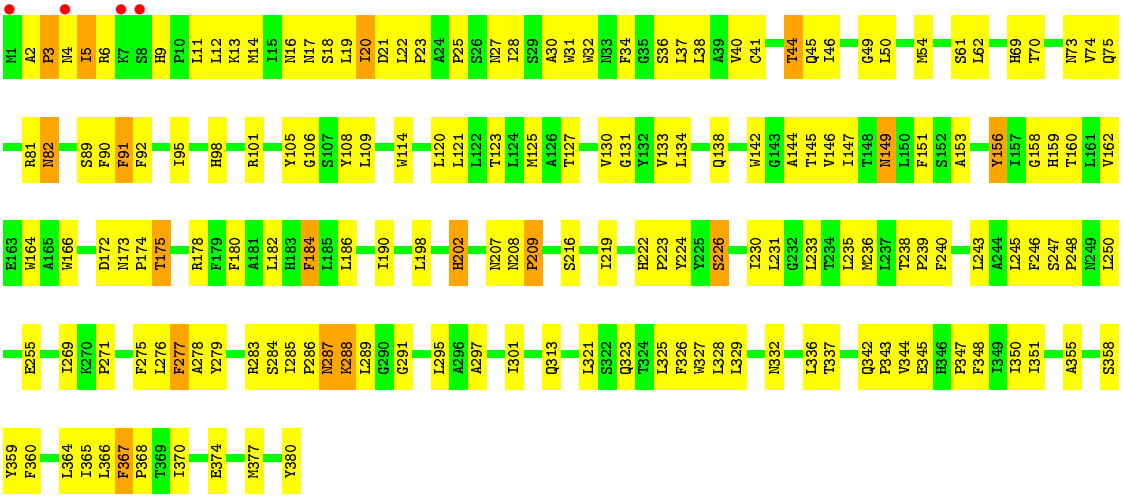
Chain O:





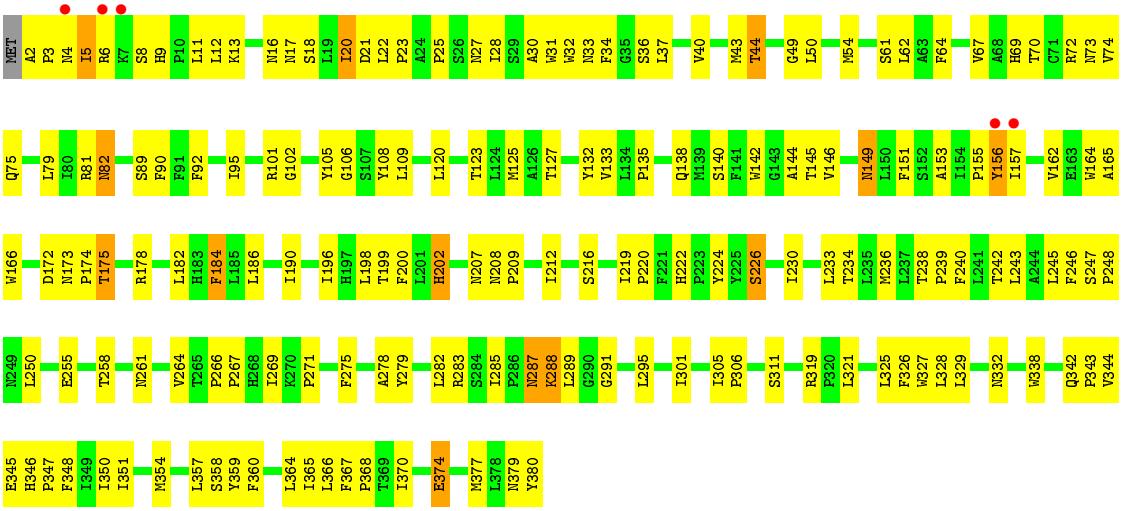
• Molecule 3: CYTOCHROME B

Chain C:



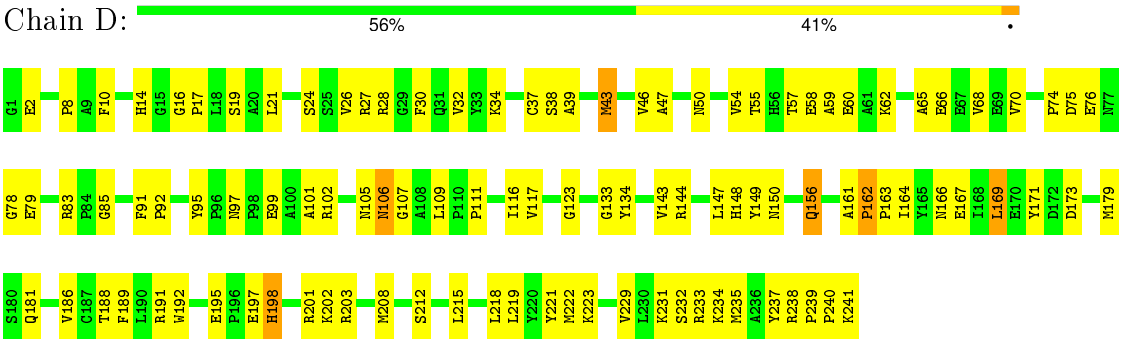
• Molecule 3: CYTOCHROME B

Chain P:



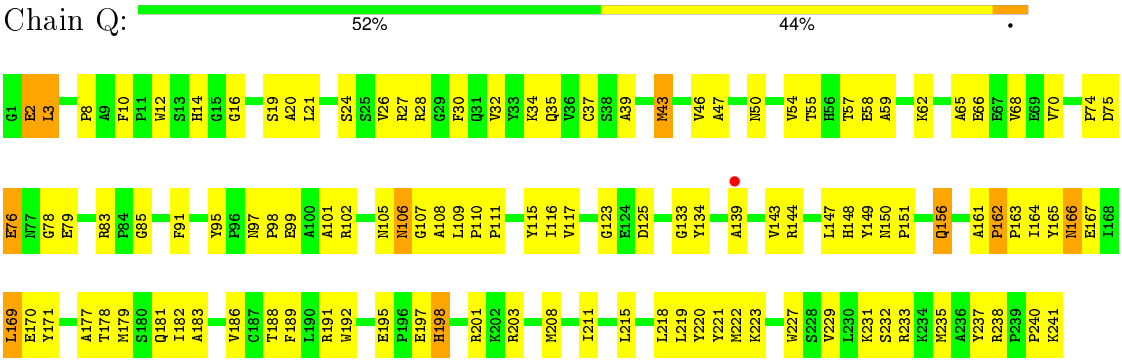
• Molecule 4: MITOCHONDRIAL CYTOCHROME C1, HEME PROTEIN

Chain D:



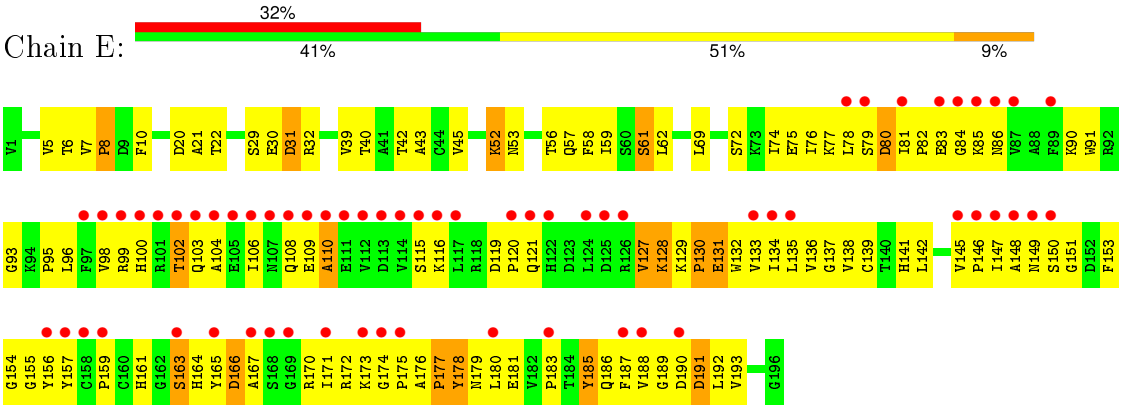
• Molecule 4: MITOCHONDRIAL CYTOCHROME C1, HEME PROTEIN

Chain Q:



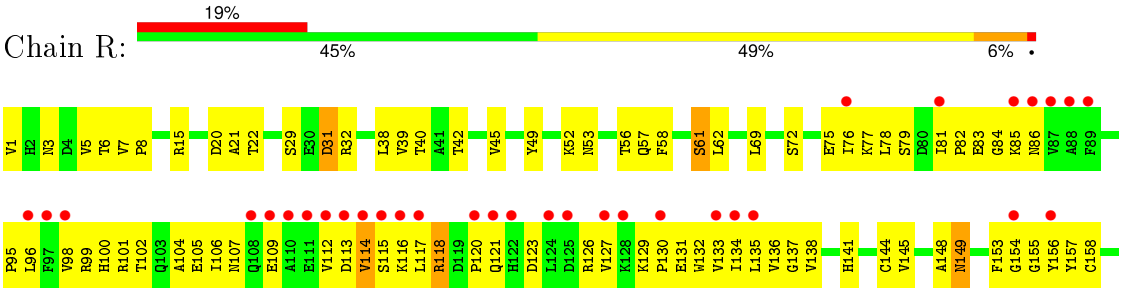
• Molecule 5: CYTOCHROME B-C1 COMPLEX SUBUNIT 5, RIESKE IRONSULFUR PROTEIN, MITOCHONDRIAL

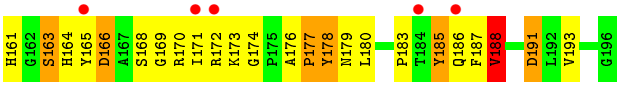
Chain E:



• Molecule 5: CYTOCHROME B-C1 COMPLEX SUBUNIT 5, RIESKE IRONSULFUR PROTEIN, MITOCHONDRIAL

Chain R:



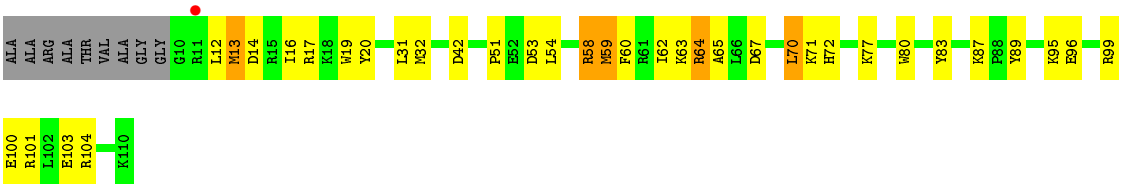


• Molecule 6: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 14 KDA PROTEIN

Chain F:

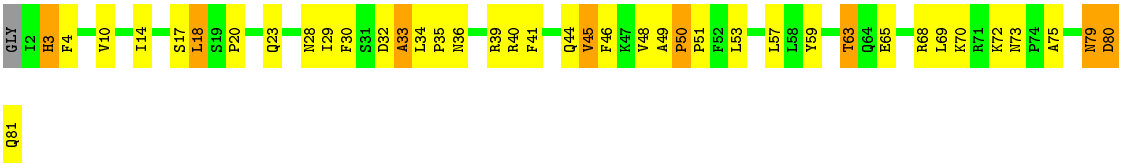
• Molecule 6: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 14 KDA PROTEIN

Chain S:



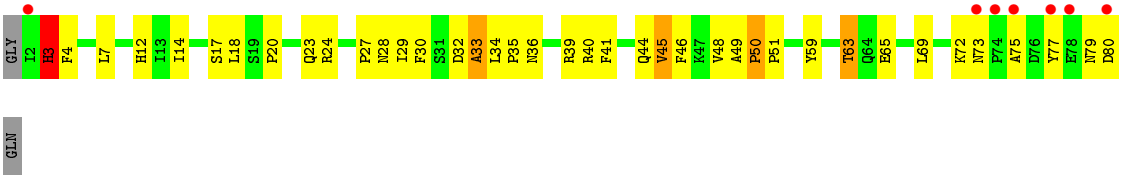
• Molecule 7: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE UBIQUINONE-BINDING PROTEIN QP-C

Chain G:



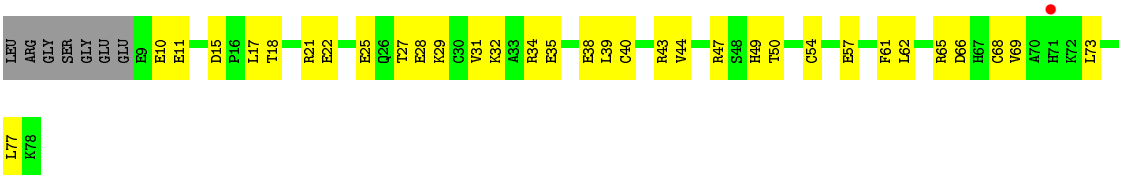
• Molecule 7: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE UBIQUINONE-BINDING PROTEIN QP-C

Chain T:



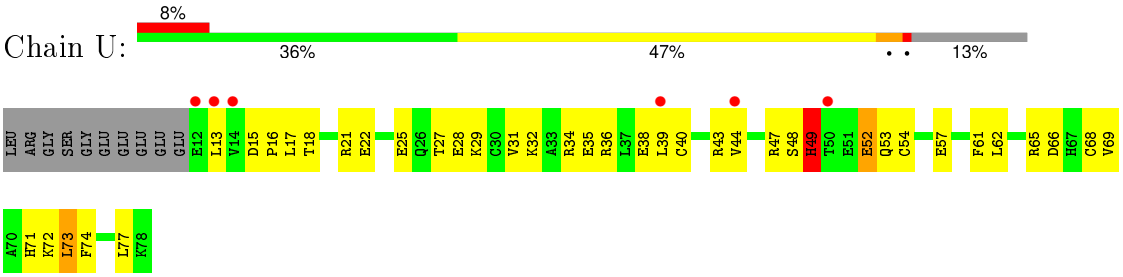
• Molecule 8: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 11 KDA PROTEIN, COMPLEX III SUBUNIT VIII

Chain H:



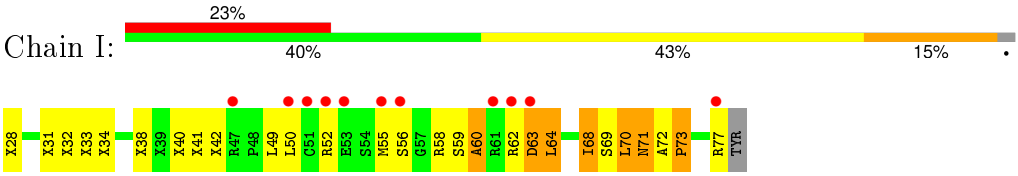
• Molecule 8: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 11 KDA PROTEIN, COMPLEX III SUBUNIT VIII

Chain U:



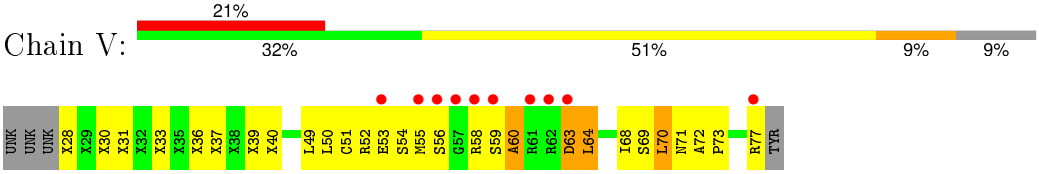
• Molecule 9: CYTOCHROME B-C1 COMPLEX SUBUNIT RIESKE, MITOCHONDRIAL

Chain I:



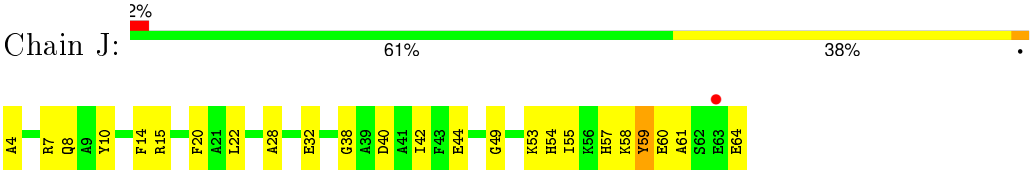
• Molecule 9: CYTOCHROME B-C1 COMPLEX SUBUNIT RIESKE, MITOCHONDRIAL

Chain V:



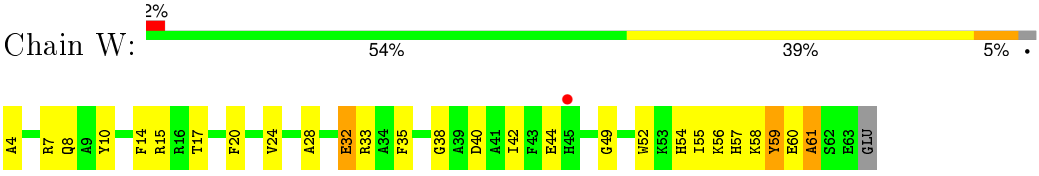
• Molecule 10: MITOCHONDRIAL UBIQUINOL-CYTOCHROME C REDUCTASE 7.2 KDA PROTEIN

Chain J:



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

Chain W:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	172.61Å 181.55Å 241.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.93 – 3.06 49.92 – 3.00	Depositor EDS
% Data completeness (in resolution range)	93.8 (28.93-3.06) 93.1 (49.92-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 3.01Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.264 , 0.294 0.253 , 0.282	Depositor DCC
R_{free} test set	2644 reflections (2.01%)	DCC
Wilson B-factor (Å^2)	75.1	Xtriage
Anisotropy	0.487	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 61.7	EDS
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Outliers	0 of 141092 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	32648	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.92% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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, IKR, CDL, UQ, FES, HEC, PEE, 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	0.44	0/3518	0.66	0/4767
1	N	0.45	0/3508	0.66	0/4753
2	B	0.40	0/3191	0.67	0/4326
2	O	0.42	0/3202	0.68	0/4343
3	C	0.51	0/3119	0.68	0/4270
3	P	0.48	0/3114	0.66	0/4263
4	D	0.46	0/1956	0.66	0/2658
4	Q	0.40	0/1956	0.63	0/2658
5	E	0.39	0/1547	0.61	0/2103
5	R	0.40	0/1543	0.62	0/2098
6	F	0.51	0/911	0.70	0/1219
6	S	0.43	0/911	0.64	0/1219
7	G	0.53	0/694	0.70	1/941 (0.1%)
7	T	0.50	0/684	0.71	1/929 (0.1%)
8	H	0.45	0/582	0.66	0/779
8	U	0.37	0/561	0.60	0/751
9	I	0.43	0/218	0.68	0/293
9	V	0.43	0/218	0.67	0/293
10	J	0.45	0/508	0.65	0/682
10	W	0.42	0/490	0.65	0/660
All	All	0.45	0/32431	0.66	2/44005 (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
6	F	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	T	18	LEU	CA-CB-CG	5.17	127.19	115.30
7	G	18	LEU	CA-CB-CG	5.11	127.05	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
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	3447	0	3362	181	0
1	N	3437	0	3349	204	0
2	B	3137	0	3131	262	0
2	O	3147	0	3146	254	0
3	C	3017	0	3063	148	0
3	P	3012	0	3058	166	0
4	D	1898	0	1846	93	0
4	Q	1898	0	1846	115	0
5	E	1513	0	1478	123	0
5	R	1509	0	1474	126	0
6	F	891	0	893	32	0
6	S	891	0	893	40	0
7	G	672	0	653	48	0
7	T	662	0	645	42	0
8	H	574	0	548	35	0
8	U	553	0	535	42	0
9	I	287	0	251	46	0
9	V	277	0	251	45	0
10	J	497	0	490	20	0
10	W	479	0	478	27	0
11	C	86	0	60	7	0
11	P	86	0	60	5	0
12	C	25	0	20	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	P	25	0	20	3	0
13	C	19	0	17	6	0
13	P	19	0	17	5	0
14	C	40	0	24	1	0
14	D	42	0	28	0	0
14	P	40	0	24	2	0
14	Q	42	0	28	2	0
15	C	70	0	85	1	0
15	E	50	0	77	0	0
15	P	54	0	72	2	0
15	R	50	0	77	1	0
16	C	6	0	8	0	0
16	P	6	0	8	2	0
17	D	43	0	30	2	0
17	Q	43	0	30	2	0
18	D	33	0	39	1	0
18	P	12	0	11	1	0
18	Q	33	0	39	1	0
19	E	4	0	0	2	0
19	R	4	0	0	2	0
20	C	8	0	0	1	0
20	E	1	0	0	0	0
20	P	8	0	0	1	0
20	R	1	0	0	0	0
All	All	32648	0	32164	1865	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 29.

The worst 5 of 1865 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)
5:E:121:GLN:HG2	5:E:170:ARG:HD3	1.26	1.14
5:E:127:VAL:HG12	5:E:128:LYS:H	1.15	1.09
2:O:76:THR:HG22	2:O:82:SER:H	1.14	1.04
3:P:328:LEU:HD12	7:T:51:PRO:HB3	1.36	1.04
2:B:76:THR:HG22	2:B:82:SER:H	1.17	1.03

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	442/446 (99%)	391 (88%)	44 (10%)	7 (2%)	12 42
1	N	440/446 (99%)	380 (86%)	49 (11%)	11 (2%)	7 30
2	B	419/441 (95%)	337 (80%)	62 (15%)	20 (5%)	3 16
2	O	420/441 (95%)	351 (84%)	57 (14%)	12 (3%)	6 27
3	C	378/380 (100%)	332 (88%)	37 (10%)	9 (2%)	7 31
3	P	377/380 (99%)	322 (85%)	47 (12%)	8 (2%)	9 35
4	D	239/241 (99%)	212 (89%)	19 (8%)	8 (3%)	5 24
4	Q	239/241 (99%)	213 (89%)	15 (6%)	11 (5%)	3 16
5	E	194/196 (99%)	142 (73%)	32 (16%)	20 (10%)	1 3
5	R	194/196 (99%)	158 (81%)	23 (12%)	13 (7%)	1 8
6	F	99/110 (90%)	89 (90%)	9 (9%)	1 (1%)	19 56
6	S	99/110 (90%)	88 (89%)	10 (10%)	1 (1%)	19 56
7	G	78/81 (96%)	63 (81%)	11 (14%)	4 (5%)	2 14
7	T	77/81 (95%)	60 (78%)	13 (17%)	4 (5%)	2 14
8	H	68/77 (88%)	57 (84%)	10 (15%)	1 (2%)	13 44
8	U	65/77 (84%)	48 (74%)	14 (22%)	3 (5%)	3 16
9	I	29/47 (62%)	20 (69%)	5 (17%)	4 (14%)	0 1
9	V	29/47 (62%)	23 (79%)	3 (10%)	3 (10%)	1 3
10	J	59/61 (97%)	46 (78%)	12 (20%)	1 (2%)	11 41
10	W	58/61 (95%)	45 (78%)	9 (16%)	4 (7%)	1 8
All	All	4003/4160 (96%)	3377 (84%)	481 (12%)	145 (4%)	4 22

5 of 145 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	21	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	22	GLU
2	B	24	LEU
2	B	26	ILE
2	B	29	LEU

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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%)	341 (93%)	24 (7%)	21 54
1	N	365/368 (99%)	341 (93%)	24 (7%)	21 54
2	B	331/347 (95%)	308 (93%)	23 (7%)	19 52
2	O	333/347 (96%)	314 (94%)	19 (6%)	25 60
3	C	328/329 (100%)	313 (95%)	15 (5%)	33 69
3	P	328/329 (100%)	314 (96%)	14 (4%)	35 71
4	D	200/200 (100%)	196 (98%)	4 (2%)	63 87
4	Q	200/200 (100%)	196 (98%)	4 (2%)	63 87
5	E	166/166 (100%)	158 (95%)	8 (5%)	31 67
5	R	165/166 (99%)	159 (96%)	6 (4%)	42 77
6	F	93/96 (97%)	90 (97%)	3 (3%)	46 79
6	S	93/96 (97%)	88 (95%)	5 (5%)	27 62
7	G	71/71 (100%)	67 (94%)	4 (6%)	26 61
7	T	70/71 (99%)	68 (97%)	2 (3%)	50 81
8	H	65/71 (92%)	64 (98%)	1 (2%)	72 90
8	U	63/71 (89%)	61 (97%)	2 (3%)	46 79
9	I	23/26 (88%)	20 (87%)	3 (13%)	5 20
9	V	23/26 (88%)	21 (91%)	2 (9%)	13 42
10	J	49/49 (100%)	46 (94%)	3 (6%)	23 57
10	W	47/49 (96%)	46 (98%)	1 (2%)	61 86

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3378/3446 (98%)	3211 (95%)	167 (5%)	31 67

5 of 167 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	F	64	ARG
1	N	86	PHE
5	R	188	VAL
7	G	3	HIS
9	I	71	ASN

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	159	GLN
5	R	186	GLN
7	G	73	ASN
1	N	10	ASN

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

29 ligands 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	$\#\|Z\| > 2$
12	IKR	C	2001	-	25,26,26	1.59	6 (24%)	31,35,35	1.21	3 (9%)
13	UQ	C	2002	-	19,19,63	2.49	11 (57%)	23,26,79	1.42	4 (17%)
14	CDL	C	2004	-	39,39,99	1.20	2 (5%)	41,51,111	1.15	4 (9%)
15	PEE	C	2007	-	48,48,50	1.31	7 (14%)	49,53,55	0.90	4 (8%)
15	PEE	C	2008	-	20,20,50	1.82	6 (30%)	21,25,55	0.65	0
16	GOL	C	2011	-	5,5,5	1.54	1 (20%)	5,5,5	0.80	0
11	HEM	C	501	3	30,50,50	2.81	9 (30%)	24,82,82	2.44	8 (33%)
11	HEM	C	502	3	30,50,50	2.45	7 (23%)	24,82,82	2.11	8 (33%)
14	CDL	D	2003	-	41,41,99	1.18	2 (4%)	43,53,111	1.10	2 (4%)
18	BOG	D	2009	-	20,20,20	1.07	1 (5%)	25,25,25	0.94	2 (8%)
18	BOG	D	2091	-	13,13,20	1.45	3 (23%)	18,18,25	1.12	2 (11%)
17	HEC	D	501	4	24,50,50	2.77	3 (12%)	19,82,82	3.53	8 (42%)
15	PEE	E	2005	-	49,49,50	1.42	9 (18%)	50,54,55	0.94	5 (10%)
19	FES	E	501	5	0,4,4	0.00	-	0,4,4	0.00	-
18	BOG	P	2010	-	12,12,20	1.52	4 (33%)	17,17,25	0.63	0
12	IKR	P	3001	-	25,26,26	1.41	4 (16%)	31,35,35	1.24	4 (12%)
13	UQ	P	3002	-	19,19,63	2.41	11 (57%)	23,26,79	1.42	4 (17%)
14	CDL	P	3004	-	39,39,99	1.22	2 (5%)	41,51,111	1.15	4 (9%)
15	PEE	P	3007	-	48,48,50	1.28	6 (12%)	49,53,55	0.87	4 (8%)
15	PEE	P	3008	-	4,4,50	3.69	4 (100%)	6,6,55	0.53	0
16	GOL	P	3011	-	5,5,5	1.30	0	5,5,5	0.65	0
11	HEM	P	501	3	30,50,50	2.87	9 (30%)	24,82,82	2.37	8 (33%)
11	HEM	P	502	3	30,50,50	2.56	9 (30%)	24,82,82	2.29	8 (33%)
14	CDL	Q	3003	-	41,41,99	1.19	3 (7%)	43,53,111	1.08	2 (4%)
18	BOG	Q	3009	-	20,20,20	1.11	3 (15%)	25,25,25	1.02	2 (8%)
18	BOG	Q	3091	-	13,13,20	1.54	3 (23%)	18,18,25	1.06	2 (11%)
17	HEC	Q	501	4	24,50,50	2.26	3 (12%)	19,82,82	3.43	7 (36%)
15	PEE	R	3005	-	49,49,50	1.50	9 (18%)	50,54,55	0.96	5 (10%)
19	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
12	IKR	C	2001	-	-	0/18/18/18	0/2/2/2
13	UQ	C	2002	-	-	0/11/35/87	0/1/1/1
14	CDL	C	2004	-	-	0/49/49/110	0/0/0/0
15	PEE	C	2007	-	-	0/52/52/54	0/0/0/0
15	PEE	C	2008	-	-	0/24/24/54	0/0/0/0
16	GOL	C	2011	-	-	0/4/4/4	0/0/0/0
11	HEM	C	501	3	-	0/10/54/54	0/0/8/8
11	HEM	C	502	3	-	0/10/54/54	0/0/8/8
14	CDL	D	2003	-	-	0/51/51/110	0/0/0/0
18	BOG	D	2009	-	-	0/11/31/31	0/1/1/1
18	BOG	D	2091	-	-	0/4/24/31	0/1/1/1
17	HEC	D	501	4	-	0/6/54/54	0/0/8/8
15	PEE	E	2005	-	-	0/53/53/54	0/0/0/0
19	FES	E	501	5	-	0/0/4/4	0/1/1/1
18	BOG	P	2010	-	-	0/2/22/31	0/1/1/1
12	IKR	P	3001	-	-	0/18/18/18	0/2/2/2
13	UQ	P	3002	-	-	0/11/35/87	0/1/1/1
14	CDL	P	3004	-	-	0/49/49/110	0/0/0/0
15	PEE	P	3007	-	-	0/52/52/54	0/0/0/0
15	PEE	P	3008	-	-	0/0/0/54	0/0/0/0
16	GOL	P	3011	-	-	0/4/4/4	0/0/0/0
11	HEM	P	501	3	-	0/10/54/54	0/0/8/8
11	HEM	P	502	3	-	0/10/54/54	0/0/8/8
14	CDL	Q	3003	-	-	0/51/51/110	0/0/0/0
18	BOG	Q	3009	-	-	0/11/31/31	0/1/1/1
18	BOG	Q	3091	-	-	0/4/24/31	0/1/1/1
17	HEC	Q	501	4	-	0/6/54/54	0/0/8/8
15	PEE	R	3005	-	-	0/53/53/54	0/0/0/0
19	FES	R	501	5	-	0/0/4/4	0/1/1/1

The worst 5 of 137 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	D	501	HEC	C3B-C2B	-9.84	1.30	1.40
17	Q	501	HEC	C3B-C2B	-8.25	1.32	1.40
17	D	501	HEC	C3C-C2C	-7.60	1.32	1.40
11	P	501	HEM	C3B-C4B	-7.03	1.45	1.51
11	C	501	HEM	C3B-C4B	-6.88	1.45	1.51

The worst 5 of 100 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
17	D	501	HEC	CBC-CAC-C3C	-10.65	103.68	127.35
17	Q	501	HEC	CBC-CAC-C3C	-10.46	104.10	127.35
17	D	501	HEC	CBB-CAB-C3B	-5.90	114.25	127.35
17	Q	501	HEC	CBB-CAB-C3B	-5.65	114.81	127.35
11	C	501	HEM	CBA-CAA-C2A	-5.57	102.54	112.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

22 monomers are involved in 52 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	C	2001	IKR	4	0
13	C	2002	UQ	6	0
14	C	2004	CDL	1	0
15	C	2007	PEE	1	0
11	C	501	HEM	3	0
11	C	502	HEM	4	0
18	D	2009	BOG	1	0
17	D	501	HEC	2	0
19	E	501	FES	2	0
18	P	2010	BOG	1	0
12	P	3001	IKR	3	0
13	P	3002	UQ	5	0
14	P	3004	CDL	2	0
15	P	3007	PEE	2	0
16	P	3011	GOL	2	0
11	P	501	HEM	2	0
11	P	502	HEM	3	0
14	Q	3003	CDL	2	0
18	Q	3009	BOG	1	0
17	Q	501	HEC	2	0
15	R	3005	PEE	1	0
19	R	501	FES	2	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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^2)	Q<0.9
1	A	444/446 (99%)	-0.20	3 (0%) 89 75	39, 74, 106, 125	0
1	N	442/446 (99%)	-0.10	4 (0%) 85 69	43, 79, 111, 119	0
2	B	421/441 (95%)	-0.05	6 (1%) 78 57	63, 90, 121, 156	0
2	O	422/441 (95%)	-0.11	8 (1%) 70 45	50, 85, 115, 134	0
3	C	380/380 (100%)	-0.37	4 (1%) 82 63	23, 45, 98, 135	0
3	P	379/380 (99%)	-0.22	5 (1%) 79 59	32, 68, 104, 128	0
4	D	241/241 (100%)	-0.38	0 100 100	37, 51, 89, 114	0
4	Q	241/241 (100%)	-0.10	1 (0%) 93 84	55, 85, 118, 130	0
5	E	196/196 (100%)	1.31	63 (32%) 1 0	39, 140, 179, 186	125 (63%)
5	R	196/196 (100%)	0.72	38 (19%) 1 0	51, 99, 146, 165	0
6	F	101/110 (91%)	-0.54	0 100 100	38, 52, 70, 90	0
6	S	101/110 (91%)	-0.17	1 (0%) 84 66	62, 84, 126, 151	0
7	G	80/81 (98%)	-0.25	0 100 100	39, 61, 118, 127	0
7	T	79/81 (97%)	0.10	7 (8%) 12 4	55, 91, 154, 161	0
8	H	70/77 (90%)	-0.36	1 (1%) 78 57	52, 74, 95, 133	0
8	U	67/77 (87%)	0.53	6 (8%) 12 4	103, 131, 152, 157	0
9	I	31/47 (65%)	1.69	11 (35%) 0 0	92, 128, 164, 167	0
9	V	31/47 (65%)	1.59	10 (32%) 1 0	87, 119, 167, 170	0
10	J	61/61 (100%)	-0.24	1 (1%) 74 52	46, 63, 107, 143	0
10	W	60/61 (98%)	0.18	1 (1%) 73 49	65, 83, 120, 129	0
All	All	4043/4160 (97%)	-0.03	170 (4%) 40 18	23, 77, 138, 186	125 (3%)

The worst 5 of 170 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	113	ASP	11.3
5	E	107	ASN	8.6
9	I	51	CYS	8.1
5	E	167	ALA	6.8
5	E	98	VAL	6.5

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^{th} 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^2)	Q<0.9
15	PEE	E	2005	50/51	0.82	0.42	6.54	79,91,105,106	0
15	PEE	P	3007	49/51	0.89	0.37	4.39	68,91,107,109	0
14	CDL	Q	3003	42/100	0.76	0.34	4.21	120,141,152,153	0
15	PEE	C	2008	21/51	0.67	0.39	3.95	132,138,141,142	0
16	GOL	C	2011	6/6	0.90	0.27	3.92	54,56,58,60	0
16	GOL	P	3011	6/6	0.90	0.32	3.73	79,79,79,80	0
14	CDL	P	3004	40/100	0.83	0.33	3.23	111,116,122,123	0
15	PEE	C	2007	49/51	0.93	0.29	3.23	49,59,88,90	0
15	PEE	R	3005	50/51	0.79	0.35	3.14	80,99,106,107	0
13	UQ	C	2002	19/63	0.88	0.28	2.76	85,89,90,90	0
18	BOG	P	2010	12/20	0.57	0.42	2.75	162,165,167,167	0
14	CDL	D	2003	42/100	0.80	0.29	2.50	112,121,125,126	0
18	BOG	Q	3009	20/20	0.90	0.32	2.40	71,99,102,102	0
13	UQ	P	3002	19/63	0.85	0.29	2.24	97,111,114,115	0
18	BOG	D	2009	20/20	0.93	0.29	2.11	51,76,80,81	0
14	CDL	C	2004	40/100	0.89	0.25	1.22	72,82,99,101	0
11	HEM	C	501	43/43	0.98	0.21	0.58	30,37,43,45	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
11	HEM	P	501	43/43	0.98	0.21	0.46	46,50,57,60	0
11	HEM	P	502	43/43	0.97	0.21	0.45	40,47,59,62	0
17	HEC	Q	501	43/43	0.96	0.23	0.21	68,71,76,78	0
12	IKR	P	3001	25/25	0.98	0.19	0.01	66,67,76,78	0
17	HEC	D	501	43/43	0.98	0.18	-0.02	38,42,46,48	0
11	HEM	C	502	43/43	0.98	0.19	-0.20	24,31,37,44	0
12	IKR	C	2001	25/25	0.99	0.16	-0.97	38,40,44,54	0
19	FES	R	501	4/4	0.98	0.09	-1.56	88,90,91,91	0
19	FES	E	501	4/4	0.95	0.11	-2.15	151,151,152,152	4
18	BOG	D	2091	13/20	0.50	0.63	-	167,171,171,171	0
15	PEE	P	3008	5/51	0.80	0.26	-	140,140,140,140	0
18	BOG	Q	3091	13/20	0.31	0.77	-	194,197,197,198	0

6.5 Other polymers (i)

There are no such residues in this entry.