



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2017 – 08:44 AM EDT

PDB ID : 3TGU  
Title : Cytochrome bc1 complex from chicken with pfvs-designed moa inhibitor bound  
Authors : Huang, L.-S.; Yang, G.-F.; Berry, E.A.  
Deposited on : unknown  
Resolution : 2.70 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 : rb-20030345  
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) : rb-20030345

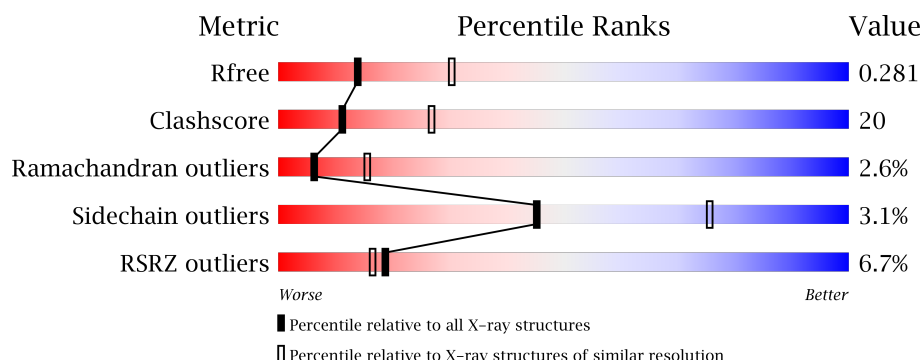
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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>0.1%</div> <div>67%</div> <div>29%</div> <div>0.1%</div> </div>
1	N	446	<div> <div>2%</div> <div>64%</div> <div>33%</div> <div>1%</div> </div>
2	B	441	<div> <div>3%</div> <div>52%</div> <div>39%</div> <div>5%</div> </div>
2	O	441	<div> <div>5%</div> <div>52%</div> <div>39%</div> <div>4%</div> </div>
3	C	380	<div> <div>3%</div> <div>74%</div> <div>26%</div> <div>0.1%</div> </div>

*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	76	
9	V	76	
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	UNL	A	501	-	-	-	X
11	UNL	N	501	-	-	-	X
14	UQ	C	504	-	-	-	X
14	UQ	P	505	-	-	-	X
15	CDL	C	505	-	-	-	X
15	CDL	Q	502	-	-	-	X
16	PEE	C	507	-	-	-	X
16	PEE	E	502	-	-	-	X
16	PEE	N	502	-	X	-	-
16	PEE	P	507	-	-	-	X
16	PEE	R	502	-	-	-	X
17	GOL	C	508	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
17	GOL	P	508	-	-	-	X
19	BOG	P	503	-	-	-	X
19	BOG	Q	503	-	-	-	X

## 2 Entry composition

There are 21 unique types of molecules in this entry. The entry contains 32733 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	421	Total	C	N	O	S	0	0	0
			3141	1974	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
			3021	2025	478	505	13			
3	P	379	Total	C	N	O	S	0	0	0
			3012	2019	477	504	12			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	FME	-	INITIATING METHIONINE	UNP P18946
P	1	FME	-	INITIATING METHIONINE	UNP P18946

- 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
			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	52	Total	C	N	O	S	0	0	2
			319	190	65	61	3			
9	V	50	Total	C	N	O	S	0	0	2
			311	186	63	59	3			

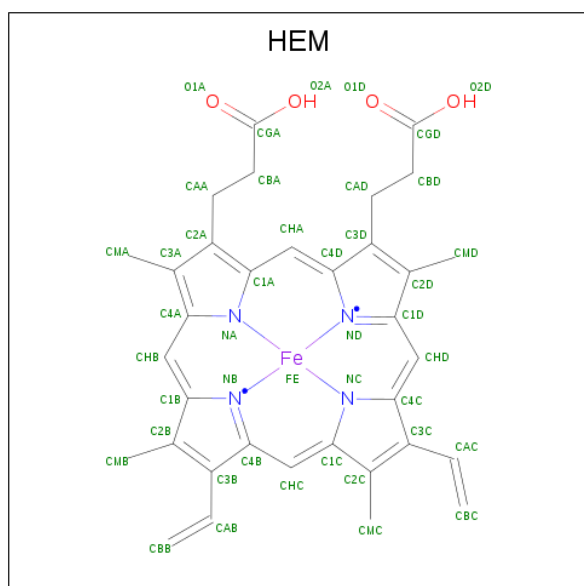
- Molecule 10 is a protein called Mitochondrial ubiquinol-cytochrome c reductase 7.2 kda protein.

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

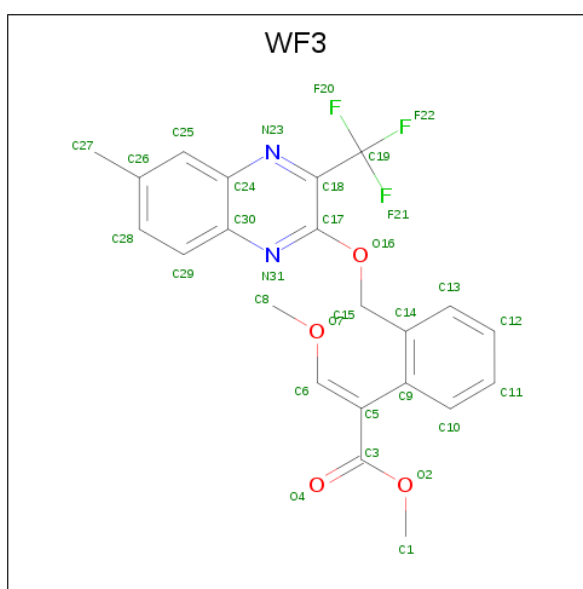
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	1	Total	O	0	0
			1	1		
11	N	1	Total	O	0	0
			1	1		

- Molecule 12 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



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

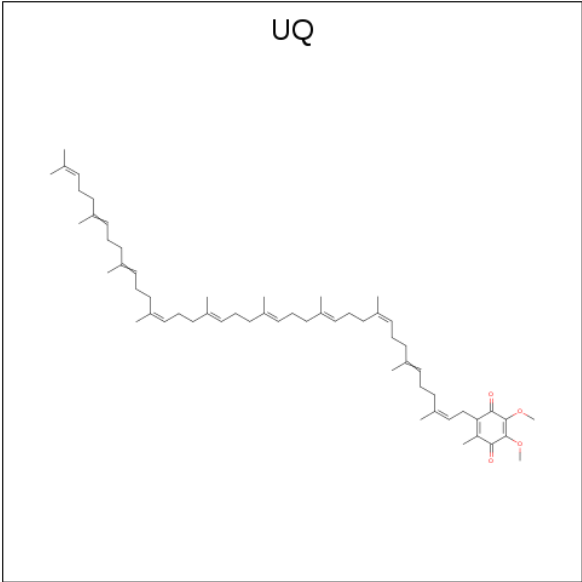
- Molecule 13 is methyl (2E)-3-methoxy-2-[2-({[6-methyl-3-(trifluoromethyl)quinoxalin-2-yl]oxy}methyl)phenyl]prop-2-enoate (three-letter code: WF3) (formula: C<sub>22</sub>H<sub>19</sub>F<sub>3</sub>N<sub>2</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
13	C	1	Total	C	F	N	O	
			31	22	3	2	4	
13	P	1	Total	C	F	N	O	
			31	22	3	2	4	

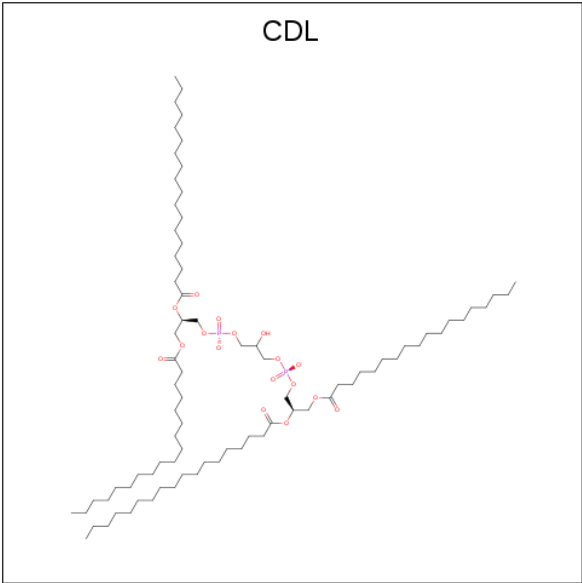
- Molecule 14 is Coenzyme Q10, (2Z,6E,10Z,14E,18E,22E,26Z)-isomer (three-letter code: UQ) (formula: C<sub>59</sub>H<sub>90</sub>O<sub>4</sub>).





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

- Molecule 15 is CARDIOLIPIN (three-letter code: CDL) (formula:  $C_{81}H_{156}O_{17}P_2$ ).



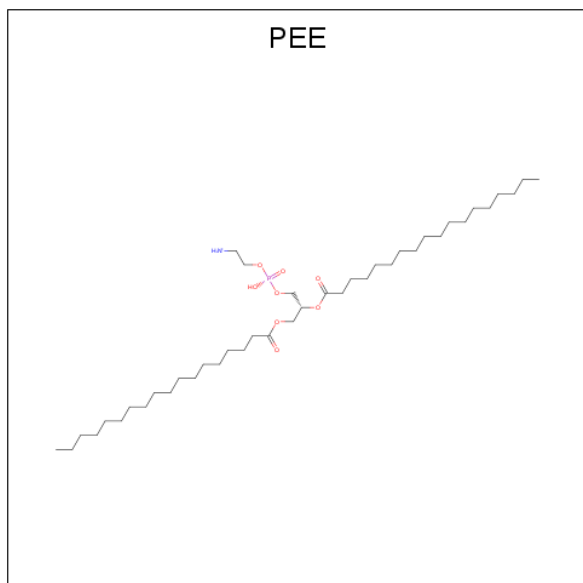
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
15	C	1	Total	C	O	P	0	0
			42	23	17	2		
15	G	1	Total	C	O	P	0	0
			40	21	17	2		

Continued on next page...

*Continued from previous page...*

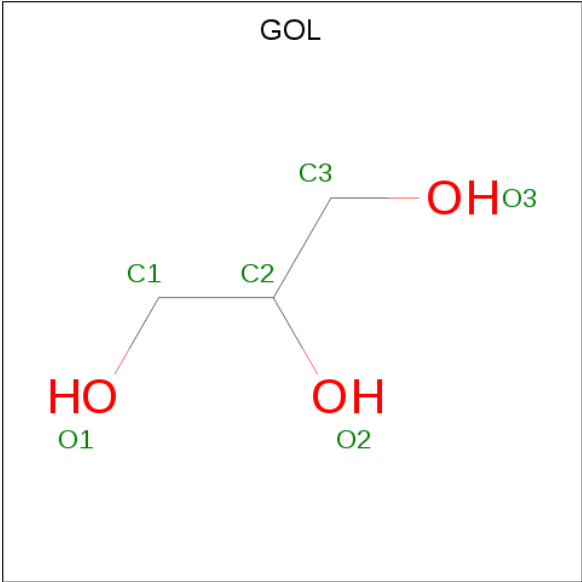
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
15	P	1	Total	C	O	P	0	0
			40	21	17	2		
15	Q	1	Total	C	O	P	0	0
			42	23	17	2		

- Molecule 16 is 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine (three-letter code: PEE) (formula:  $C_{41}H_{83}NO_8P$ ).



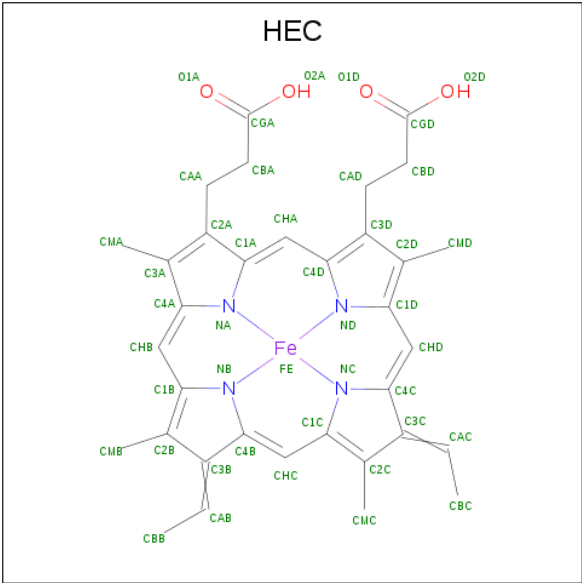
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
16	C	1	Total	C	N	O	P	0	0
			49	39	1	8	1		
16	C	1	Total	C	O	P		0	0
			21	12	8	1			
16	E	1	Total	C	N	O	P	0	0
			50	40	1	8	1		
16	N	1	Total	O	P			0	0
			5	4	1				
16	P	1	Total	C	N	O	P	0	0
			49	39	1	8	1		
16	R	1	Total	C	N	O	P	0	0
			49	39	1	8	1		

- Molecule 17 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



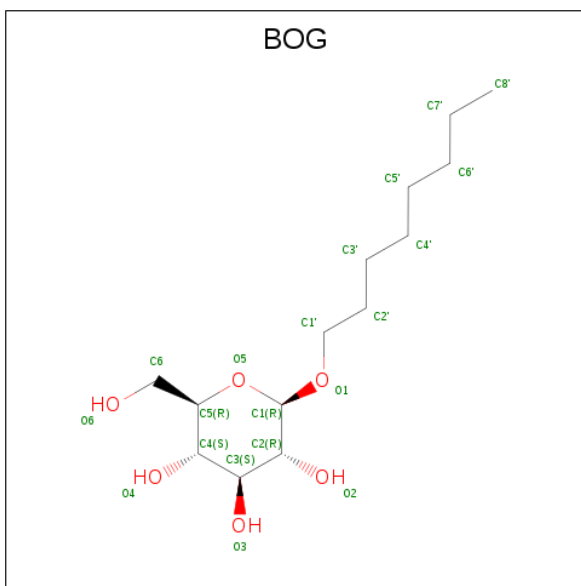
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
17	C	1	Total	C	O	0	0
			6	3	3		
17	P	1	Total	C	O	0	0
			6	3	3		

- Molecule 18 is HEME C (three-letter code: HEC) (formula:  $C_{34}H_{34}FeN_4O_4$ ).



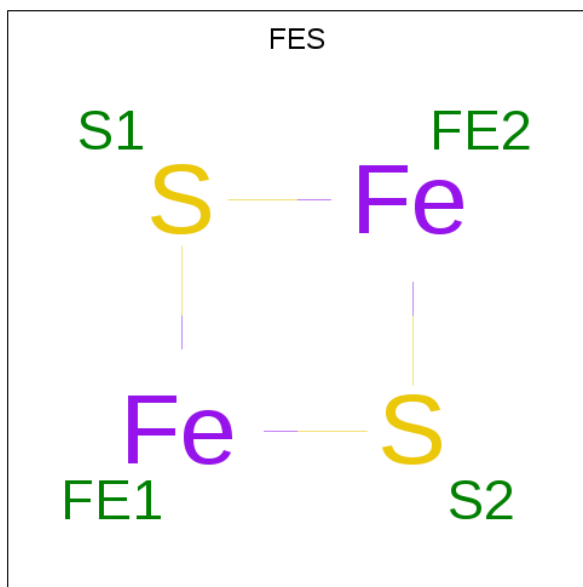
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
18	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
18	Q	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 19 is B-OCTYLGLUCOSIDE (three-letter code: BOG) (formula:  $C_{14}H_{28}O_6$ ).



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

- Molecule 20 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula:  $Fe_2S_2$ ).



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

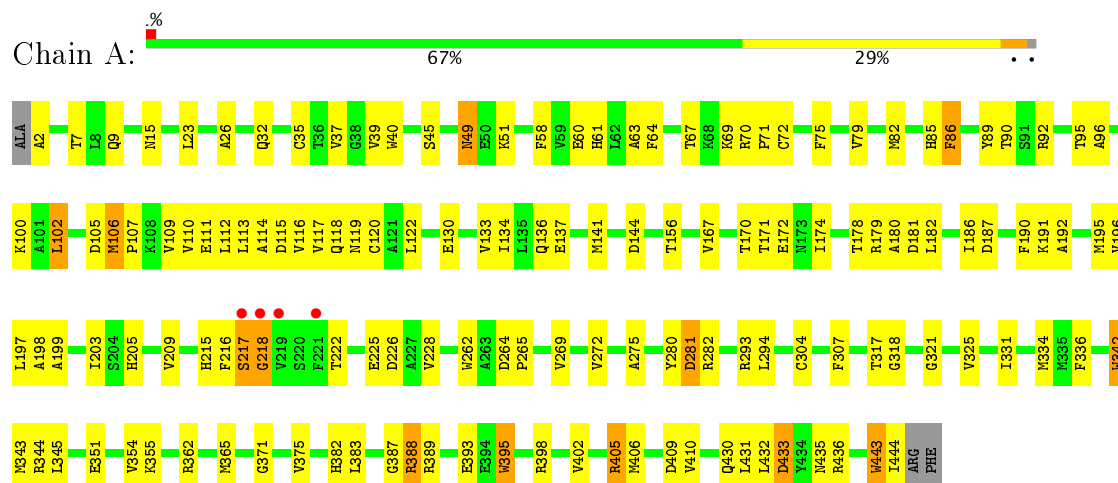
- Molecule 21 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
21	C	9	Total	O	0	0
			9	9		
21	E	1	Total	O	0	0
			1	1		
21	P	10	Total	O	0	0
			10	10		
21	R	1	Total	O	0	0
			1	1		

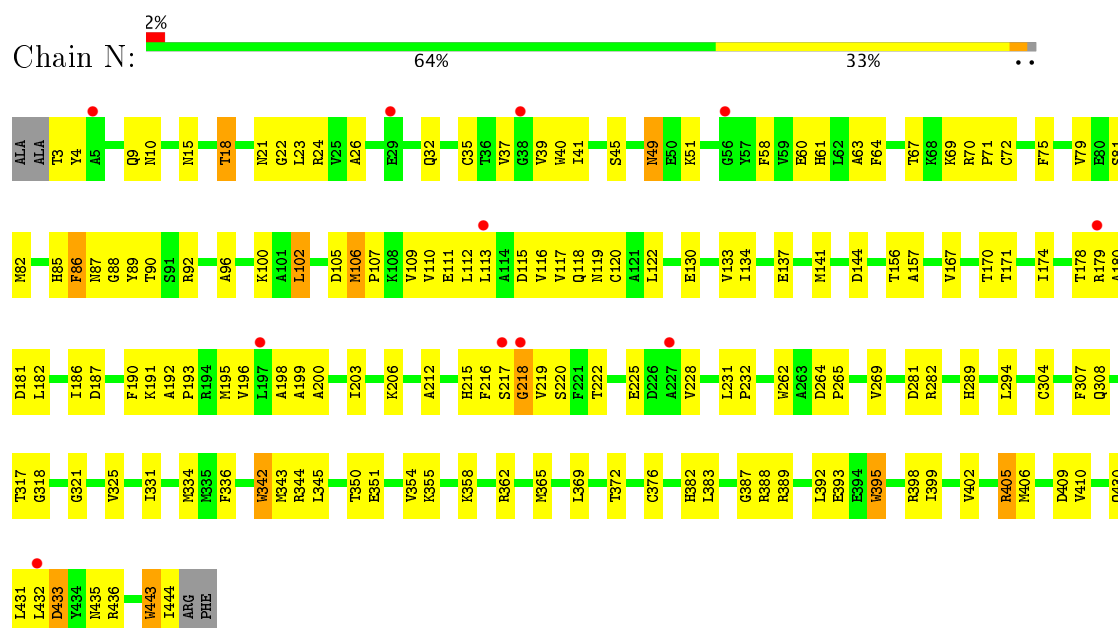
### 3 Residue-property plots

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.

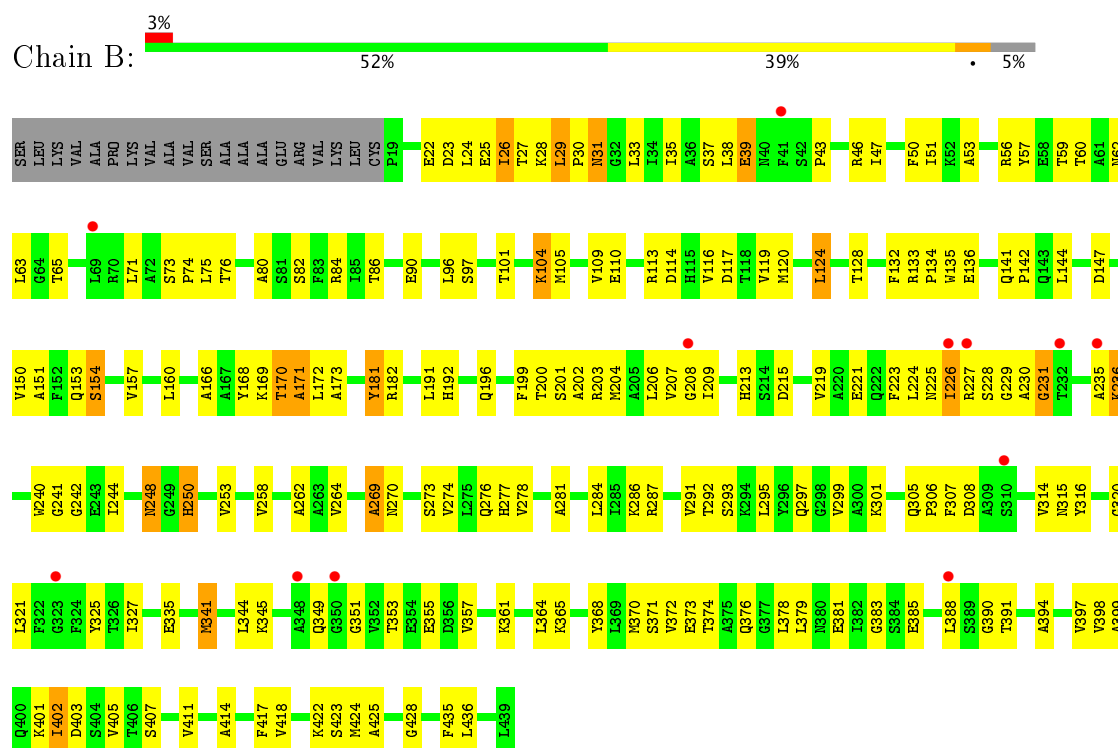
- Molecule 1: Mitochondrial ubiquinol-cytochrome-c reductase complex core protein i



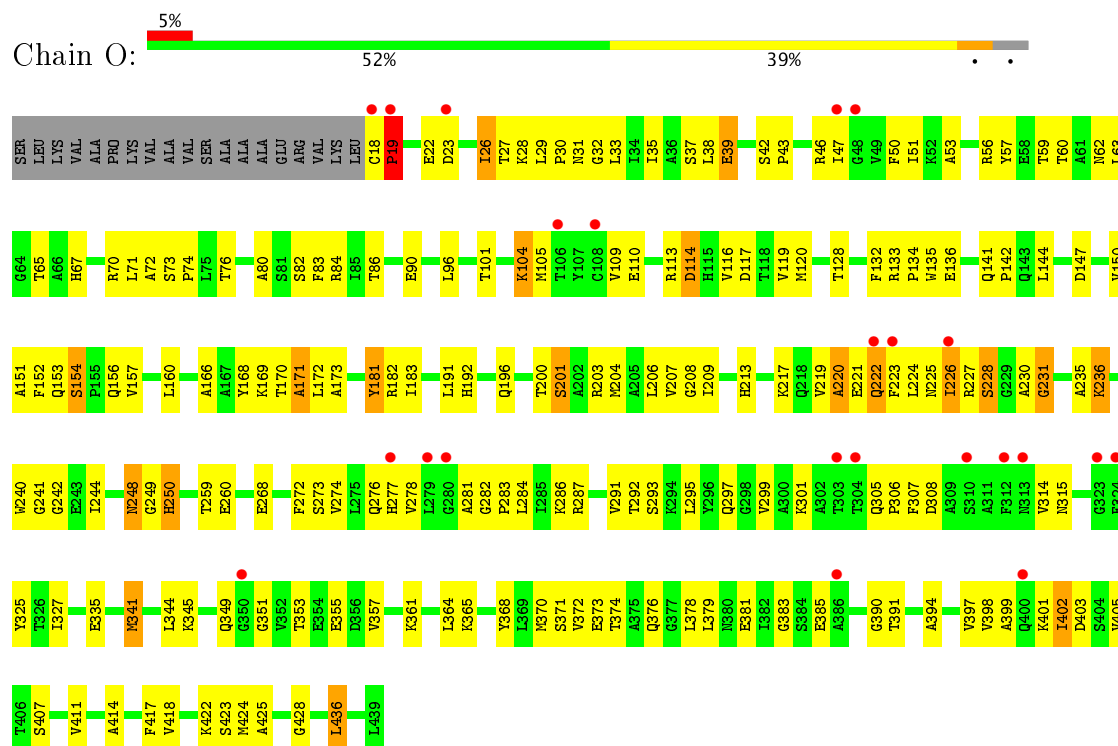
- Molecule 1: Mitochondrial ubiquinol-cytochrome-c reductase complex core protein i



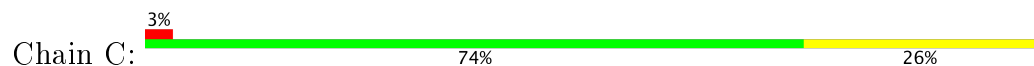
- Molecule 2: Mitochondrial ubiquinol-cytochrome-c reductase complex core protein 2

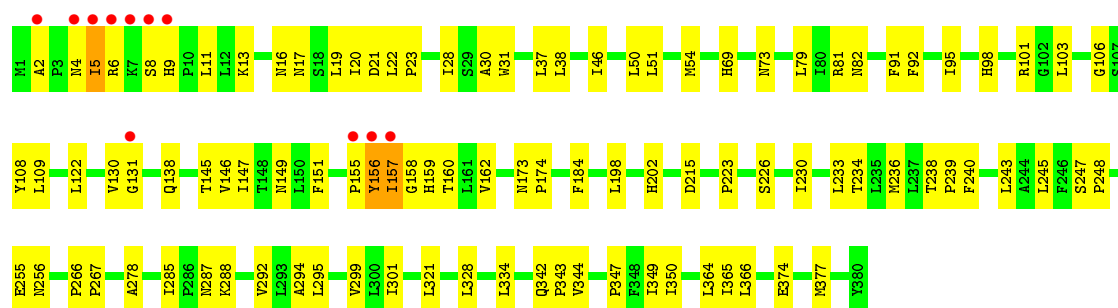


• Molecule 2: Mitochondrial ubiquinol-cytochrome-c reductase complex core protein 2

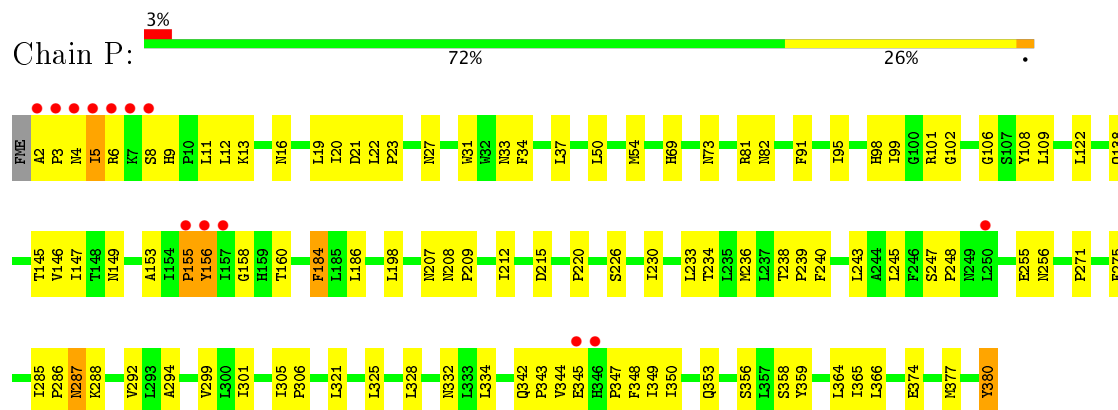


• Molecule 3: Cytochrome b

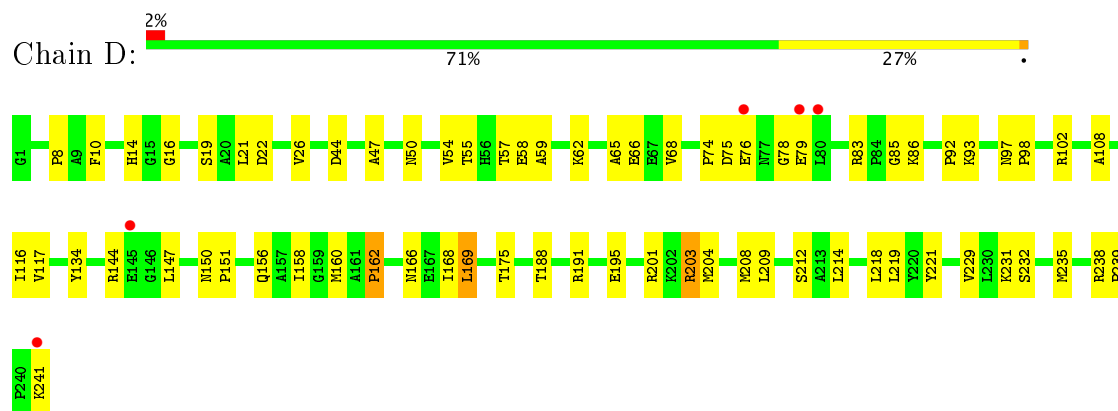




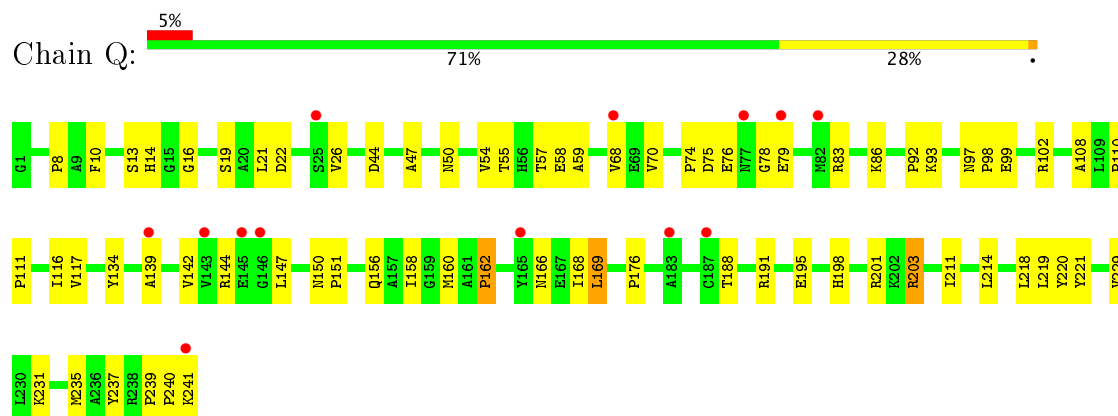
• Molecule 3: Cytochrome b



• Molecule 4: Mitochondrial cytochrome c1, heme protein

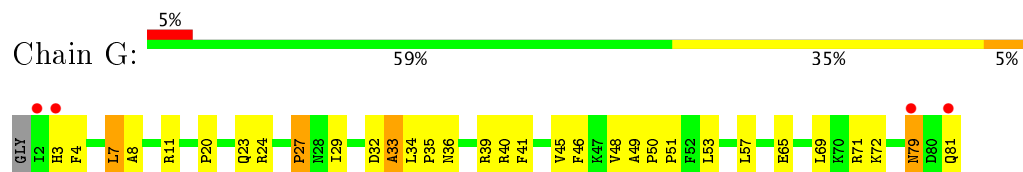


• Molecule 4: Mitochondrial cytochrome c1, heme protein

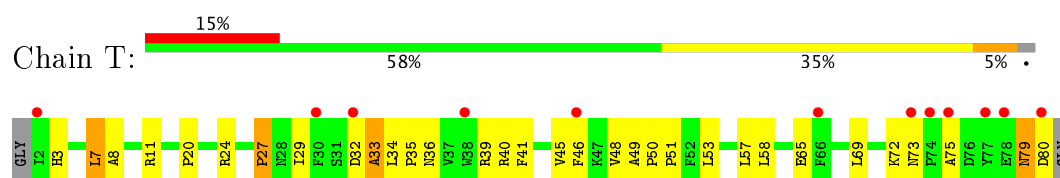




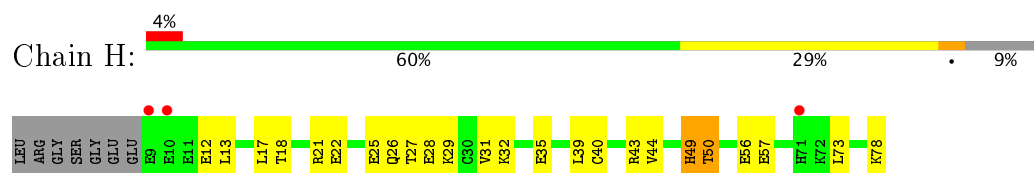
- Chain E:



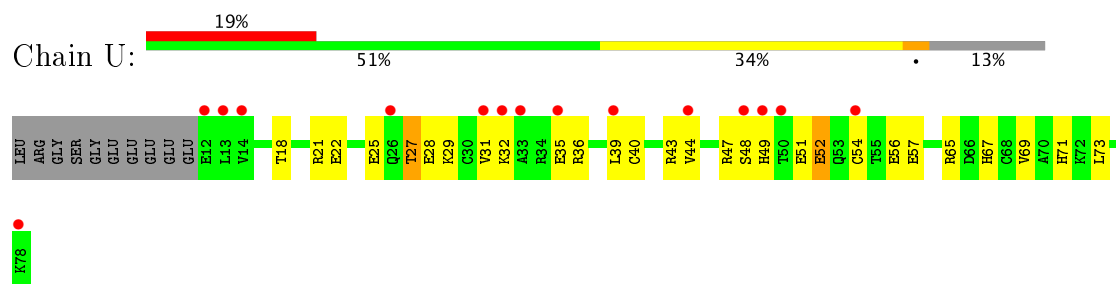
- Molecule 7: Mitochondrial ubiquinol-cytochrome c reductase ubiquinone-binding protein qp-c



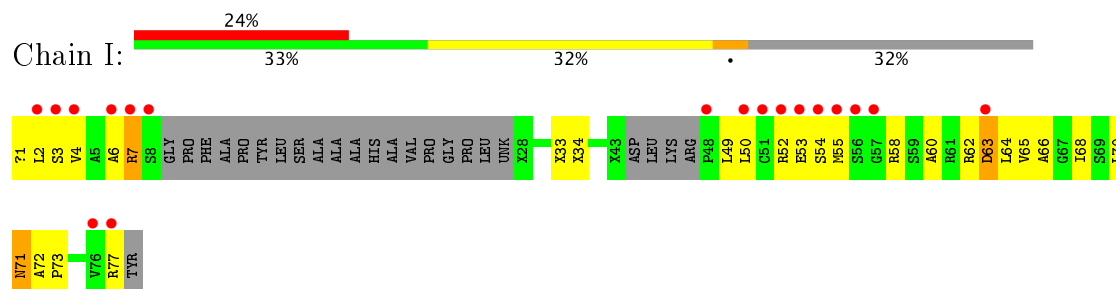
- Molecule 8: Mitochondrial ubiquinol-cytochrome c reductase 11 kda protein, complex iii subunit viii



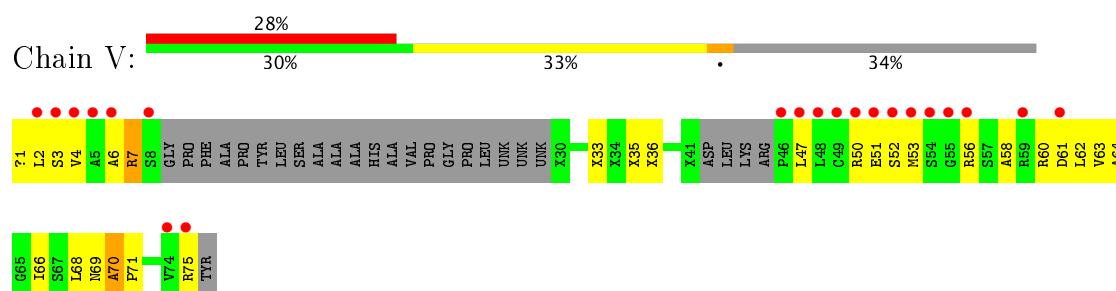
- Molecule 8: Mitochondrial ubiquinol-cytochrome c reductase 11 kda protein, complex iii subunit viii



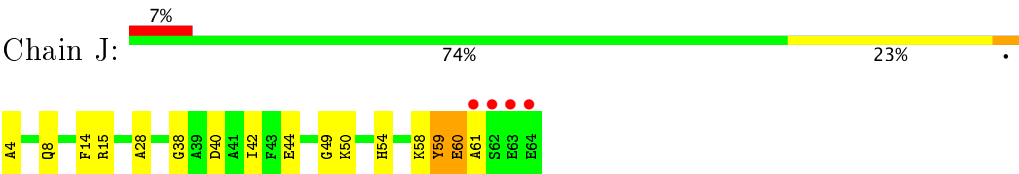
- Molecule 9: Cytochrome b-c1 complex subunit Rieske, mitochondrial



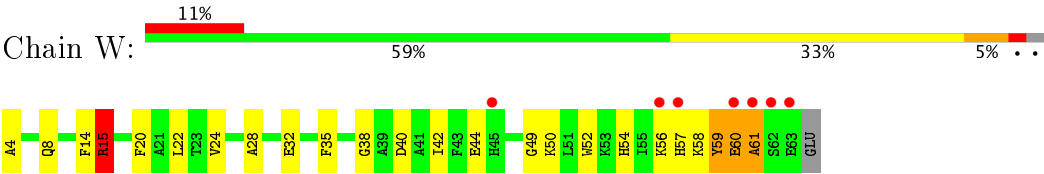
- Molecule 9: Cytochrome b-c1 complex subunit Rieske, mitochondrial



- 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, $\alpha$ , $\beta$ , $\gamma$	172.68Å 183.31Å 241.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.98 – 2.70 24.98 – 2.70	Depositor EDS
% Data completeness (in resolution range)	91.1 (24.98-2.70) 91.2 (24.98-2.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 2.72Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.258 , 0.289 0.253 , 0.281	Depositor DCC
$R_{free}$ test set	3737 reflections (1.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	60.9	Xtriage
Anisotropy	0.687	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 75.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	32733	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.85% 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, CDL, UQ, AME, UNL, FES, HEC, WF3, FME, 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	0.39	0/3513	0.64	0/4760
1	N	0.40	0/3508	0.63	0/4753
2	B	0.35	0/3196	0.62	0/4334
2	O	0.37	0/3202	0.63	0/4343
3	C	0.45	0/3114	0.65	0/4263
3	P	0.42	0/3114	0.63	0/4263
4	D	0.41	0/1956	0.62	0/2658
4	Q	0.37	0/1956	0.60	0/2658
5	E	0.35	0/1547	0.59	0/2103
5	R	0.35	0/1543	0.60	0/2098
6	F	0.47	0/911	0.65	0/1219
6	S	0.38	0/911	0.60	0/1219
7	G	0.44	0/694	0.65	0/941
7	T	0.41	0/684	0.64	0/929
8	H	0.39	0/582	0.60	0/779
8	U	0.32	0/561	0.58	0/751
9	I	0.36	0/251	0.62	0/336
9	V	0.34	0/251	0.60	0/336
10	J	0.39	0/508	0.59	0/682
10	W	0.41	0/490	0.58	0/660
All	All	0.39	0/32492	0.62	0/44085

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 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	119	0
1	N	3437	0	3349	134	0
2	B	3141	0	3142	195	0
2	O	3147	0	3146	213	0
3	C	3021	0	3068	74	0
3	P	3012	0	3058	79	0
4	D	1898	0	1846	58	0
4	Q	1898	0	1846	59	0
5	E	1513	0	1478	108	0
5	R	1509	0	1474	92	0
6	F	891	0	893	16	0
6	S	891	0	893	29	0
7	G	672	0	653	29	0
7	T	662	0	645	33	0
8	H	574	0	548	18	0
8	U	553	0	535	29	0
9	I	319	0	281	43	0
9	V	311	0	283	48	0
10	J	497	0	490	13	0
10	W	479	0	478	22	0
11	A	1	0	0	0	0
11	N	1	0	0	0	0
12	C	86	0	60	5	0
12	P	86	0	60	4	0
13	C	31	0	19	1	0
13	P	31	0	19	1	0
14	C	19	0	17	3	0
14	P	19	0	17	2	0
15	C	42	0	28	1	0
15	G	40	0	24	1	0
15	P	40	0	24	3	0
15	Q	42	0	28	1	0
16	C	70	0	85	2	0
16	E	50	0	77	0	0
16	N	5	0	0	0	0
16	P	49	0	72	2	0
16	R	49	0	71	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	C	6	0	8	0	0
17	P	6	0	8	0	0
18	D	43	0	30	2	0
18	Q	43	0	30	1	0
19	D	33	0	39	0	0
19	P	12	0	11	0	0
19	Q	33	0	39	0	0
20	E	4	0	0	0	0
20	R	4	0	0	0	0
21	C	9	0	0	0	0
21	E	1	0	0	0	0
21	P	10	0	0	2	0
21	R	1	0	0	0	0
All	All	32733	0	32226	1293	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 20.

The worst 5 of 1293 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:O:76:THR:HG22	2:O:82:SER:H	1.05	1.19
1:N:178:THR:HG22	1:N:180:ALA:H	1.10	1.14
1:A:178:THR:HG22	1:A:180:ALA:H	1.11	1.13
2:B:76:THR:HG22	2:B:82:SER:H	1.12	1.12
1:A:69:LYS:HD2	1:A:70:ARG:HH21	1.12	1.10

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%)	408 (92%)	25 (6%)	8 (2%)	10	25
1	N	440/446 (99%)	409 (93%)	24 (6%)	7 (2%)	11	28
2	B	419/441 (95%)	351 (84%)	52 (12%)	16 (4%)	4	8
2	O	420/441 (95%)	359 (86%)	46 (11%)	15 (4%)	4	9
3	C	378/380 (100%)	350 (93%)	24 (6%)	4 (1%)	17	40
3	P	377/380 (99%)	348 (92%)	23 (6%)	6 (2%)	11	28
4	D	239/241 (99%)	222 (93%)	16 (7%)	1 (0%)	38	66
4	Q	239/241 (99%)	221 (92%)	15 (6%)	3 (1%)	14	35
5	E	194/196 (99%)	147 (76%)	31 (16%)	16 (8%)	1	1
5	R	194/196 (99%)	151 (78%)	32 (16%)	11 (6%)	2	3
6	F	99/110 (90%)	95 (96%)	4 (4%)	0	100	100
6	S	99/110 (90%)	92 (93%)	7 (7%)	0	100	100
7	G	78/81 (96%)	65 (83%)	11 (14%)	2 (3%)	6	15
7	T	77/81 (95%)	66 (86%)	9 (12%)	2 (3%)	6	15
8	H	68/77 (88%)	63 (93%)	4 (6%)	1 (2%)	12	30
8	U	65/77 (84%)	55 (85%)	8 (12%)	2 (3%)	5	11
9	I	34/76 (45%)	24 (71%)	6 (18%)	4 (12%)	0	0
9	V	34/76 (45%)	24 (71%)	6 (18%)	4 (12%)	0	0
10	J	59/61 (97%)	54 (92%)	5 (8%)	0	100	100
10	W	58/61 (95%)	52 (90%)	4 (7%)	2 (3%)	4	10
All	All	4012/4218 (95%)	3556 (89%)	352 (9%)	104 (3%)	6	15

5 of 104 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	218	GLY
1	A	433	ASP
2	B	26	ILE
2	B	39	GLU
2	B	226	ILE

### 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%)	353 (97%)	12 (3%)	43	73
1	N	365/368 (99%)	351 (96%)	14 (4%)	38	68
2	B	332/347 (96%)	317 (96%)	15 (4%)	32	62
2	O	333/347 (96%)	320 (96%)	13 (4%)	37	68
3	C	328/328 (100%)	318 (97%)	10 (3%)	46	76
3	P	328/328 (100%)	318 (97%)	10 (3%)	46	76
4	D	200/200 (100%)	198 (99%)	2 (1%)	80	93
4	Q	200/200 (100%)	198 (99%)	2 (1%)	80	93
5	E	166/166 (100%)	158 (95%)	8 (5%)	30	59
5	R	165/166 (99%)	161 (98%)	4 (2%)	54	83
6	F	93/96 (97%)	91 (98%)	2 (2%)	57	84
6	S	93/96 (97%)	91 (98%)	2 (2%)	57	84
7	G	71/71 (100%)	69 (97%)	2 (3%)	49	79
7	T	70/71 (99%)	68 (97%)	2 (3%)	48	77
8	H	65/71 (92%)	64 (98%)	1 (2%)	70	90
8	U	63/71 (89%)	63 (100%)	0	100	100
9	I	26/45 (58%)	25 (96%)	1 (4%)	38	68
9	V	26/45 (58%)	26 (100%)	0	100	100
10	J	49/49 (100%)	47 (96%)	2 (4%)	35	66
10	W	47/49 (96%)	44 (94%)	3 (6%)	20	45
All	All	3385/3482 (97%)	3280 (97%)	105 (3%)	45	75

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	F	64	ARG
1	N	58	PHE
5	R	190	ASP
6	F	70	LEU
10	J	59	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 74 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	73	ASN
2	O	31	ASN
7	T	23	GLN
7	G	79	ASN
1	N	274	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
3	FME	C	1	3	8,8,10	2.02	2 (25%)	7,8,11	1.29	1 (14%)

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
3	FME	C	1	3	-	0/4/8/11	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1	FME	CN-N	-4.81	1.33	1.46
3	C	1	FME	CA-C	2.65	1.53	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1	FME	CN-N-CA	2.51	121.79	113.60

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 31 ligands modelled in this entry, 2 are unknown - 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
12	HEM	C	501	3	28,50,50	1.89	6 (21%)	17,82,82	1.72	4 (23%)
12	HEM	C	502	3	28,50,50	1.89	8 (28%)	17,82,82	1.98	4 (23%)
13	WF3	C	503	-	32,33,33	2.78	14 (43%)	39,47,47	2.36	13 (33%)
14	UQ	C	504	-	19,19,63	2.51	9 (47%)	23,26,79	1.01	1 (4%)
15	CDL	C	505	-	41,41,99	1.19	2 (4%)	43,53,111	1.11	5 (11%)
16	PEE	C	506	-	48,48,50	1.28	7 (14%)	50,53,55	0.89	4 (8%)
16	PEE	C	507	-	20,20,50	1.76	7 (35%)	22,25,55	0.64	0
17	GOL	C	508	-	5,5,5	1.38	0	5,5,5	0.67	0
18	HEC	D	501	4	28,50,50	2.05	6 (21%)	16,82,82	1.55	4 (25%)
19	BOG	D	502	-	20,20,20	0.88	0	25,25,25	0.86	1 (4%)
19	BOG	D	503	-	13,13,20	1.41	3 (23%)	18,18,25	1.17	2 (11%)
20	FES	E	501	5	0,4,4	0.00	-	0,4,4	0.00	-
16	PEE	E	502	-	49,49,50	1.45	10 (20%)	51,54,55	1.02	5 (9%)
15	CDL	G	101	-	39,39,99	1.20	2 (5%)	41,51,111	1.17	4 (9%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
16	PEE	N	502	-	4,4,50	3.49	4 (100%)	6,6,55	0.50	0
12	HEM	P	501	3	28,50,50	1.88	6 (21%)	17,82,82	1.56	4 (23%)
12	HEM	P	502	3	28,50,50	1.85	6 (21%)	17,82,82	1.77	3 (17%)
19	BOG	P	503	-	12,12,20	1.49	3 (25%)	17,17,25	0.54	0
13	WF3	P	504	-	32,33,33	2.92	14 (43%)	39,47,47	2.37	13 (33%)
14	UQ	P	505	-	19,19,63	2.49	10 (52%)	23,26,79	0.99	1 (4%)
15	CDL	P	506	-	39,39,99	1.22	2 (5%)	41,51,111	1.18	4 (9%)
16	PEE	P	507	-	48,48,50	1.26	6 (12%)	50,53,55	0.87	4 (8%)
17	GOL	P	508	-	5,5,5	1.41	1 (20%)	5,5,5	0.67	0
18	HEC	Q	501	4	28,50,50	1.81	2 (7%)	16,82,82	1.58	4 (25%)
15	CDL	Q	502	-	41,41,99	1.19	1 (2%)	43,53,111	1.12	5 (11%)
19	BOG	Q	503	-	20,20,20	0.96	2 (10%)	25,25,25	0.88	1 (4%)
19	BOG	Q	504	-	13,13,20	1.32	2 (15%)	18,18,25	1.09	2 (11%)
20	FES	R	501	5	0,4,4	0.00	-	0,4,4	0.00	-
16	PEE	R	502	-	47,47,50	1.36	8 (17%)	48,51,55	0.77	1 (2%)

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	HEM	C	501	3	-	0/6/54/54	0/0/8/8
12	HEM	C	502	3	-	0/6/54/54	0/0/8/8
13	WF3	C	503	-	-	0/24/24/24	0/3/3/3
14	UQ	C	504	-	-	0/11/35/87	0/1/1/1
15	CDL	C	505	-	-	0/51/51/110	0/0/0/0
16	PEE	C	506	-	-	0/52/52/54	0/0/0/0
16	PEE	C	507	-	-	0/24/24/54	0/0/0/0
17	GOL	C	508	-	-	0/4/4/4	0/0/0/0
18	HEC	D	501	4	-	0/6/54/54	0/0/8/8
19	BOG	D	502	-	-	0/11/31/31	0/1/1/1
19	BOG	D	503	-	-	0/4/24/31	0/1/1/1
20	FES	E	501	5	-	0/0/4/4	0/1/1/1
16	PEE	E	502	-	-	0/53/53/54	0/0/0/0
15	CDL	G	101	-	-	0/49/49/110	0/0/0/0
16	PEE	N	502	-	-	0/0/0/54	0/0/0/0
12	HEM	P	501	3	-	0/6/54/54	0/0/8/8
12	HEM	P	502	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
19	BOG	P	503	-	-	0/2/22/31	0/1/1/1
13	WF3	P	504	-	-	0/24/24/24	0/3/3/3
14	UQ	P	505	-	-	0/11/35/87	0/1/1/1
15	CDL	P	506	-	-	0/49/49/110	0/0/0/0
16	PEE	P	507	-	-	0/52/52/54	0/0/0/0
17	GOL	P	508	-	-	0/4/4/4	0/0/0/0
18	HEC	Q	501	4	-	0/6/54/54	0/0/8/8
15	CDL	Q	502	-	-	0/51/51/110	0/0/0/0
19	BOG	Q	503	-	-	0/11/31/31	0/1/1/1
19	BOG	Q	504	-	-	0/4/24/31	0/1/1/1
20	FES	R	501	5	-	0/0/4/4	0/1/1/1
16	PEE	R	502	-	-	0/49/49/54	0/0/0/0

The worst 5 of 141 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	Q	501	HEC	C3B-C2B	-6.43	1.34	1.40
18	D	501	HEC	C3C-C2C	-6.13	1.34	1.40
13	P	504	WF3	C9-C5	-5.67	1.42	1.49
18	D	501	HEC	C3B-C2B	-5.43	1.35	1.40
18	Q	501	HEC	C3C-C2C	-5.38	1.35	1.40

The worst 5 of 89 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	P	504	WF3	C8-O7-C6	-5.73	108.57	115.71
13	C	503	WF3	C8-O7-C6	-5.23	109.20	115.71
12	C	502	HEM	C4C-C3C-C2C	-5.00	103.41	106.90
13	C	503	WF3	O2-C3-O4	-3.65	116.42	123.54
13	C	503	WF3	C26-C25-C24	-3.61	118.09	121.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

17 monomers are involved in 30 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	C	501	HEM	1	0
12	C	502	HEM	4	0
13	C	503	WF3	1	0
14	C	504	UQ	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	C	505	CDL	1	0
16	C	506	PEE	2	0
18	D	501	HEC	2	0
15	G	101	CDL	1	0
12	P	501	HEM	2	0
12	P	502	HEM	2	0
13	P	504	WF3	1	0
14	P	505	UQ	2	0
15	P	506	CDL	3	0
16	P	507	PEE	2	0
18	Q	501	HEC	1	0
15	Q	502	CDL	1	0
16	R	502	PEE	1	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%)	-0.12	4 (0%) 84 85	40, 70, 101, 123	0
1	N	442/446 (99%)	0.07	11 (2%) 58 58	44, 76, 104, 117	0
2	B	421/441 (95%)	0.16	12 (2%) 52 52	55, 88, 120, 147	0
2	O	422/441 (95%)	0.18	23 (5%) 26 24	49, 81, 114, 133	0
3	C	379/380 (99%)	-0.19	11 (2%) 52 52	29, 47, 92, 135	0
3	P	379/380 (99%)	-0.05	13 (3%) 46 45	37, 68, 106, 127	0
4	D	241/241 (100%)	-0.23	5 (2%) 64 65	37, 49, 89, 118	0
4	Q	241/241 (100%)	0.41	13 (5%) 26 25	51, 82, 115, 128	0
5	E	196/196 (100%)	1.36	55 (28%) 1 0	41, 118, 173, 190	0
5	R	196/196 (100%)	0.73	36 (18%) 1 1	49, 99, 152, 173	0
6	F	101/110 (91%)	-0.45	0 100 100	31, 51, 71, 100	0
6	S	101/110 (91%)	0.22	6 (5%) 23 21	60, 79, 123, 146	0
7	G	80/81 (98%)	0.09	4 (5%) 30 28	37, 62, 116, 128	0
7	T	79/81 (97%)	0.78	12 (15%) 2 2	55, 91, 150, 165	0
8	H	70/77 (90%)	-0.02	3 (4%) 36 34	45, 69, 101, 145	0
8	U	67/77 (87%)	1.20	15 (22%) 1 1	97, 129, 149, 155	0
9	I	37/76 (48%)	2.33	18 (48%) 0 0	80, 124, 163, 166	0
9	V	37/76 (48%)	2.43	21 (56%) 0 0	61, 126, 162, 163	0
10	J	61/61 (100%)	-0.09	4 (6%) 19 17	46, 65, 107, 170	0
10	W	60/61 (98%)	0.26	7 (11%) 5 4	62, 81, 128, 138	0
All	All	4053/4218 (96%)	0.20	273 (6%) 19 16	29, 74, 134, 190	0

The worst 5 of 273 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	113	ASP	12.3
8	H	9	GLU	9.0
5	E	104	ALA	7.8
5	E	108	GLN	7.6
8	U	13	LEU	7.4

## 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
3	FME	C	1	9/11	0.64	0.62	-	139,141,148,153	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
14	UQ	P	505	19/63	0.76	0.41	8.37	80,127,140,146	0
19	BOG	P	503	12/20	0.58	0.39	7.49	126,144,149,158	0
15	CDL	Q	502	42/100	0.56	0.45	6.98	79,169,202,203	0
11	UNL	N	501	1/-	0.95	0.68	6.70	47,47,47,47	0
15	CDL	C	505	42/100	0.70	0.35	6.09	65,134,159,163	0
14	UQ	C	504	19/63	0.82	0.32	5.42	73,96,104,108	0
16	PEE	E	502	50/51	0.82	0.32	5.25	59,86,114,117	0
11	UNL	A	501	1/-	0.90	0.48	4.53	48,48,48,48	0
17	GOL	P	508	6/6	0.92	0.25	4.23	67,76,82,84	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
17	GOL	C	508	6/6	0.90	0.26	3.83	64,72,75,76	0
16	PEE	R	502	49/51	0.77	0.31	3.52	51,98,121,125	0
16	PEE	P	507	49/51	0.87	0.31	2.70	72,91,119,123	0
16	PEE	C	507	21/51	0.73	0.30	2.54	64,135,154,156	0
19	BOG	Q	503	20/20	0.91	0.27	2.09	72,100,105,113	0
15	CDL	P	506	40/100	0.83	0.27	1.94	110,125,138,141	0
16	PEE	C	506	49/51	0.92	0.23	1.94	33,60,91,103	0
15	CDL	G	101	40/100	0.92	0.21	1.46	46,83,125,131	0
12	HEM	P	501	43/43	0.98	0.22	0.92	39,58,68,81	0
12	HEM	C	502	43/43	0.98	0.21	0.91	21,39,48,59	0
12	HEM	C	501	43/43	0.98	0.21	0.81	23,41,51,66	0
13	WF3	C	503	31/31	0.97	0.17	0.37	30,40,53,60	0
13	WF3	P	504	31/31	0.96	0.18	0.24	51,64,73,78	0
12	HEM	P	502	43/43	0.99	0.17	-0.08	36,50,69,74	0
19	BOG	D	502	20/20	0.95	0.14	-0.20	44,64,72,76	0
18	HEC	Q	501	43/43	0.97	0.15	-0.39	48,60,77,85	0
18	HEC	D	501	43/43	0.99	0.12	-0.53	7,31,49,55	0
20	FES	R	501	4/4	0.99	0.08	-1.41	78,88,91,94	0
20	FES	E	501	4/4	0.97	0.08	-2.11	125,129,129,130	0
16	PEE	N	502	5/51	0.83	0.20	-	132,134,139,140	0
19	BOG	Q	504	13/20	0.54	0.50	-	179,186,190,192	0
19	BOG	D	503	13/20	0.49	0.56	-	140,156,162,164	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.