



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 28, 2016 – 10:25 PM EST

PDB ID : 5KLI  
Title : Rhodobacter sphaeroides bc1 with stigmatellin and antimycin  
Authors : Xia, D.; Esser, L.; Zhou, F.; Tang, W.K.; Yu, C.A.  
Deposited on : 2016-06-24  
Resolution : 3.00 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028320  
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) : rb-20028320

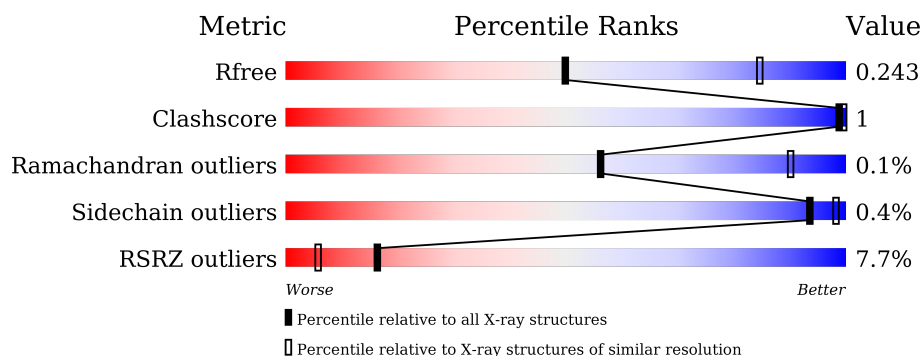
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	445	
1	E	445	
1	K	445	
1	O	445	
2	B	272	
2	F	272	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	L	272	
2	P	272	
3	C	187	
3	G	187	
3	M	187	
3	Q	187	

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
8	LOP	E	1005	-	-	-	X
8	LOP	O	1005	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 55093 atoms, of which 27155 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Cytochrome b.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	428	Total	C	H	N	O	S	0	0	0
			6841	2319	3406	545	556	15			
1	E	428	Total	C	H	N	O	S	0	0	0
			6841	2319	3406	545	556	15			
1	K	428	Total	C	H	N	O	S	0	0	0
			6841	2319	3406	545	556	15			
1	O	428	Total	C	H	N	O	S	0	0	0
			6841	2319	3406	545	556	15			

- Molecule 2 is a protein called Cytochrome c1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	256	Total	C	H	N	O	S	0	0	0
			3792	1240	1839	326	374	13			
2	F	256	Total	C	H	N	O	S	0	0	0
			3792	1240	1839	326	374	13			
2	L	256	Total	C	H	N	O	S	0	0	0
			3791	1240	1838	326	374	13			
2	P	256	Total	C	H	N	O	S	0	0	0
			3792	1240	1839	326	374	13			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	98	PRO	ALA	variant	UNP Q02760
B	264	GLY	-	expression tag	UNP Q02760
B	265	THR	-	expression tag	UNP Q02760
B	266	GLY	-	expression tag	UNP Q02760
B	267	HIS	-	expression tag	UNP Q02760
B	268	HIS	-	expression tag	UNP Q02760
B	269	HIS	-	expression tag	UNP Q02760
B	270	HIS	-	expression tag	UNP Q02760

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	271	HIS	-	expression tag	UNP Q02760
B	272	HIS	-	expression tag	UNP Q02760
F	98	PRO	ALA	variant	UNP Q02760
F	264	GLY	-	expression tag	UNP Q02760
F	265	THR	-	expression tag	UNP Q02760
F	266	GLY	-	expression tag	UNP Q02760
F	267	HIS	-	expression tag	UNP Q02760
F	268	HIS	-	expression tag	UNP Q02760
F	269	HIS	-	expression tag	UNP Q02760
F	270	HIS	-	expression tag	UNP Q02760
F	271	HIS	-	expression tag	UNP Q02760
F	272	HIS	-	expression tag	UNP Q02760
L	98	PRO	ALA	variant	UNP Q02760
L	264	GLY	-	expression tag	UNP Q02760
L	265	THR	-	expression tag	UNP Q02760
L	266	GLY	-	expression tag	UNP Q02760
L	267	HIS	-	expression tag	UNP Q02760
L	268	HIS	-	expression tag	UNP Q02760
L	269	HIS	-	expression tag	UNP Q02760
L	270	HIS	-	expression tag	UNP Q02760
L	271	HIS	-	expression tag	UNP Q02760
L	272	HIS	-	expression tag	UNP Q02760
P	98	PRO	ALA	variant	UNP Q02760
P	264	GLY	-	expression tag	UNP Q02760
P	265	THR	-	expression tag	UNP Q02760
P	266	GLY	-	expression tag	UNP Q02760
P	267	HIS	-	expression tag	UNP Q02760
P	268	HIS	-	expression tag	UNP Q02760
P	269	HIS	-	expression tag	UNP Q02760
P	270	HIS	-	expression tag	UNP Q02760
P	271	HIS	-	expression tag	UNP Q02760
P	272	HIS	-	expression tag	UNP Q02760

- Molecule 3 is a protein called Ubiquinol-cytochrome c reductase iron-sulfur subunit.

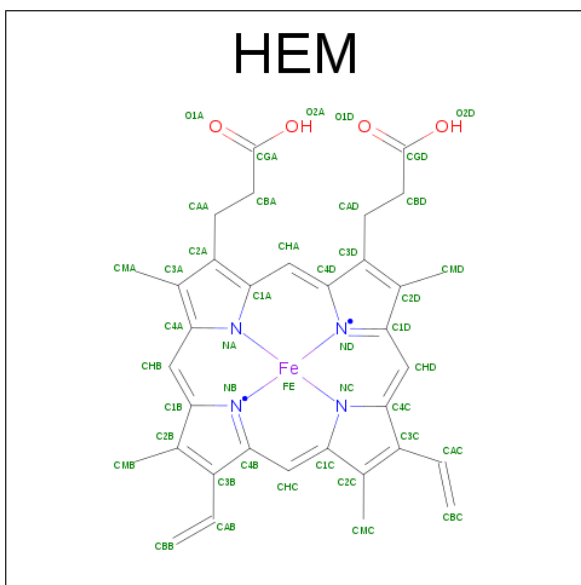
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	179	Total	C	H	N	O	S	0	0	0
			2645	845	1304	237	253	6			
3	G	179	Total	C	H	N	O	S	0	0	0
			2645	845	1304	237	253	6			
3	M	179	Total	C	H	N	O	S	0	0	0
			2645	845	1304	237	253	6			

*Continued on next page...*

Continued from previous page...

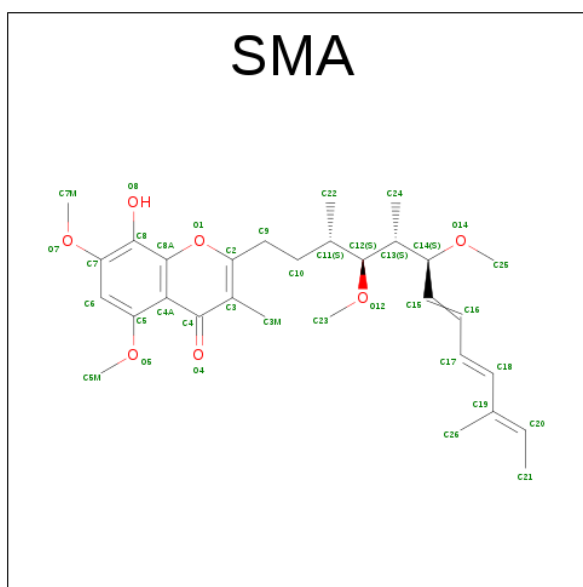
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	Q	179	Total	C	H	N	O	S	0	0	0
			2645	845	1304	237	253	6			

- Molecule 4 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



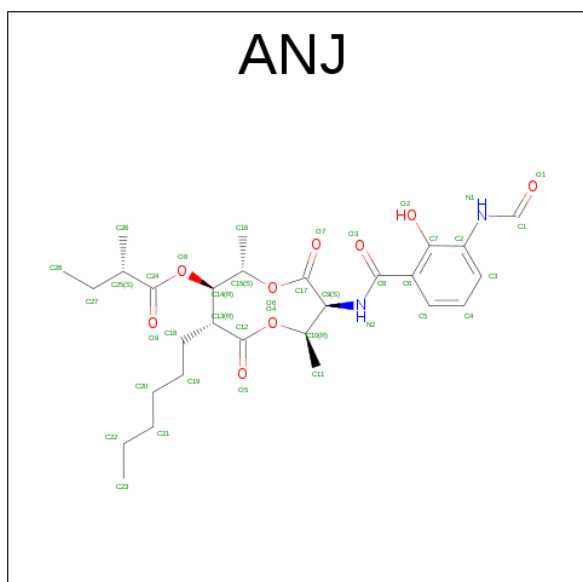
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
4	A	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
4	E	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
4	E	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
4	K	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
4	K	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
4	O	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		
4	O	1	Total	C	Fe	H	N	O	0	0
			73	34	1	30	4	4		

- Molecule 5 is STIGMATELLIN A (three-letter code: SMA) (formula:  $C_{30}H_{42}O_7$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			79	30	42	7		
5	E	1	Total	C	H	O	0	0
			79	30	42	7		
5	K	1	Total	C	H	O	0	0
			79	30	42	7		
5	O	1	Total	C	H	O	0	0
			79	30	42	7		

- Molecule 6 is (2R,3S,6S,7R,8R)-3-{[3-(FORMYLAMINO)-2-HYDROXYBENZOYL]AMINO O}-8-HEXYL-2,6-DIMETHYL-4,9-DIOXO-1,5-DIOXONAN-7-YL (2S)-2-METHYLBUTANOATE (three-letter code: ANJ) (formula: C<sub>28</sub>H<sub>40</sub>N<sub>2</sub>O<sub>9</sub>).

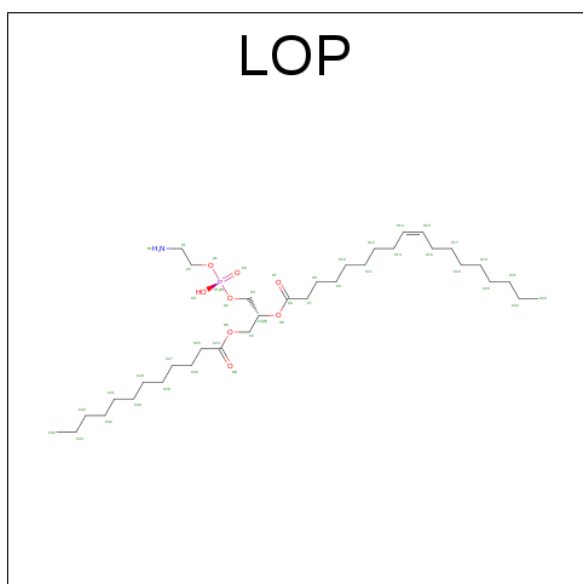


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	H	N	O	0	0
			78	28	39	2	9		
6	E	1	Total	C	H	N	O	0	0
			78	28	39	2	9		
6	K	1	Total	C	H	N	O	0	0
			78	28	39	2	9		
6	O	1	Total	C	H	N	O	0	0
			78	28	39	2	9		

- Molecule 7 is STRONTIUM ION (three-letter code: SR) (formula: Sr).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	P	1	Total	Sr	0	0
			1	1		
7	K	1	Total	Sr	0	0
			1	1		
7	B	1	Total	Sr	0	0
			1	1		
7	A	1	Total	Sr	0	0
			1	1		
7	L	1	Total	Sr	0	0
			1	1		
7	F	1	Total	Sr	0	0
			1	1		

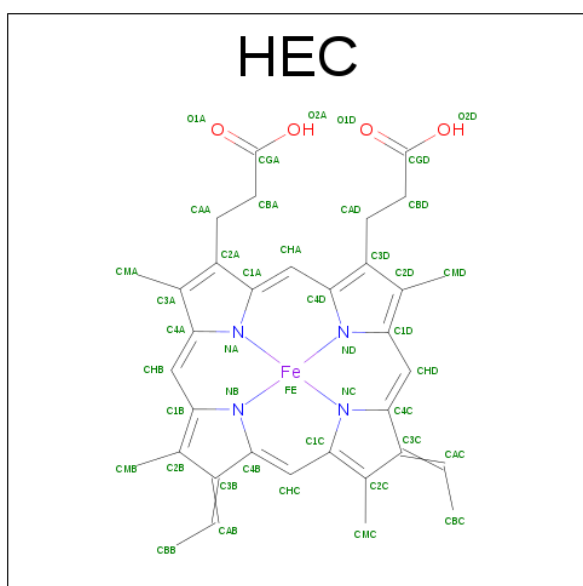
- Molecule 8 is (1R)-2-{[(R)-(2-AMINOETHOXY)(HYDROXY)PHOSPHORYL]OXY}-1-[(DODECANOYLOXY)METHYL]ETHYL (9Z)-OCTADEC-9-ENOATE (three-letter code: LOP) (formula: C<sub>35</sub>H<sub>68</sub>NO<sub>8</sub>P).





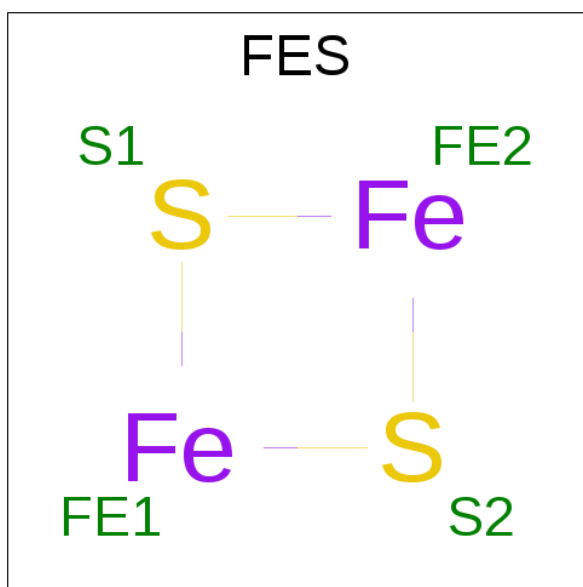
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	A	1	Total	C	H	N	O	P	
			112	35	67	1	8	1	
8	E	1	Total	C	H	N	O	P	
			112	35	67	1	8	1	
8	K	1	Total	C	H	N	O	P	
			112	35	67	1	8	1	
8	O	1	Total	C	H	N	O	P	
			112	35	67	1	8	1	

- Molecule 9 is HEME C (three-letter code: HEC) (formula:  $C_{34}H_{34}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	B	1	Total	C	Fe	H	N	O	
			75	34	1	32	4	4	
9	F	1	Total	C	Fe	H	N	O	
			75	34	1	32	4	4	
9	L	1	Total	C	Fe	H	N	O	
			75	34	1	32	4	4	
9	P	1	Total	C	Fe	H	N	O	
			75	34	1	32	4	4	

- Molecule 10 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula:  $Fe_2S_2$ ).

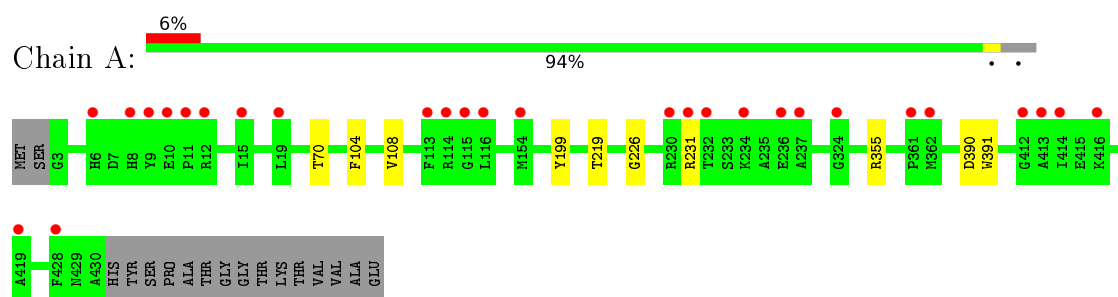


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	C	1	Total	Fe	S	0	0
			4	2	2		
10	G	1	Total	Fe	S	0	0
			4	2	2		
10	M	1	Total	Fe	S	0	0
			4	2	2		
10	Q	1	Total	Fe	S	0	0
			4	2	2		

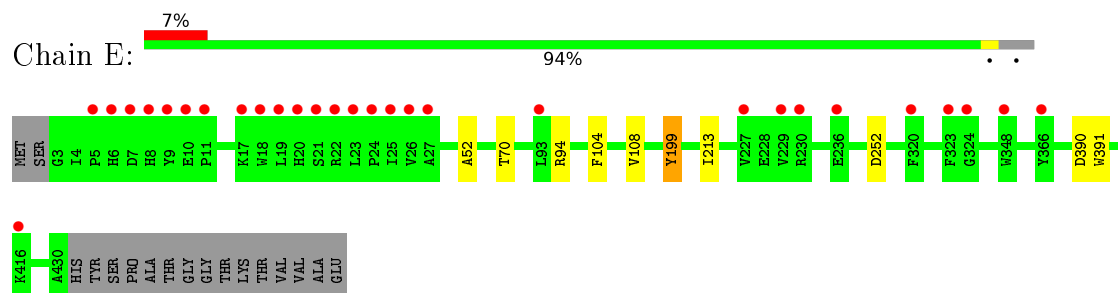
### 3 Residue-property plots [i](#)

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.

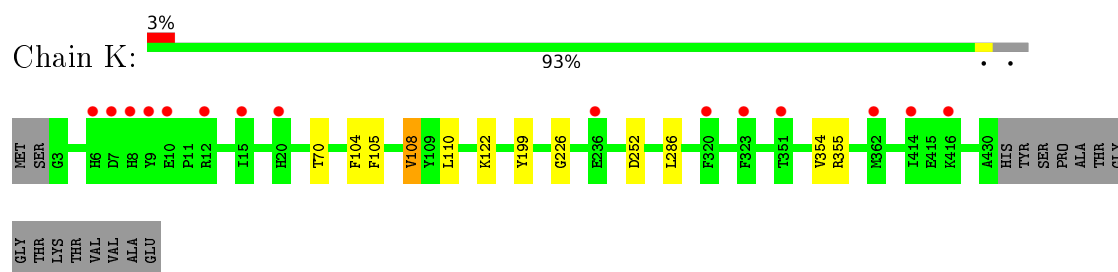
- Molecule 1: Cytochrome b



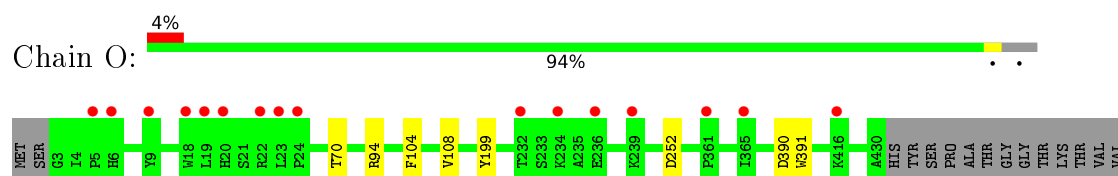
- Molecule 1: Cytochrome b



- Molecule 1: Cytochrome b

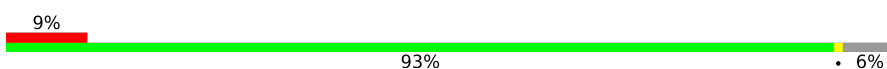


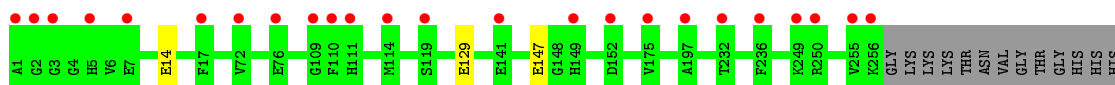
- Molecule 1: Cytochrome b



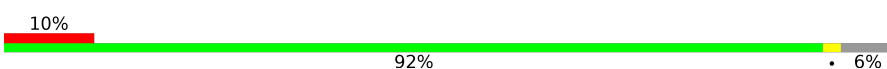
ALA  
GLU

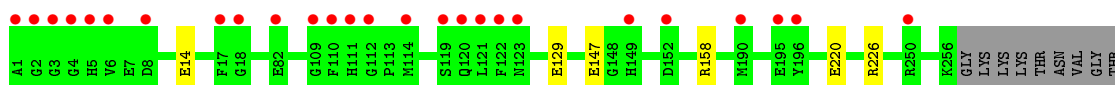
- Molecule 2: Cytochrome c1

Chain B: 

HIS  
HIS  
HIS

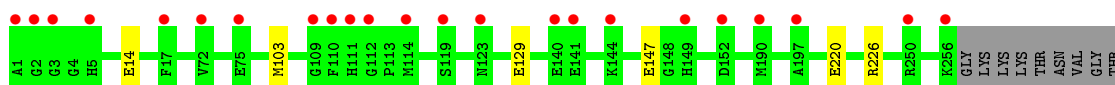
- Molecule 2: Cytochrome c1

Chain F: 

GLY  
HIS  
HIS  
HIS  
HIS  
HIS  
HIS

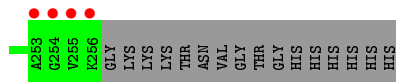
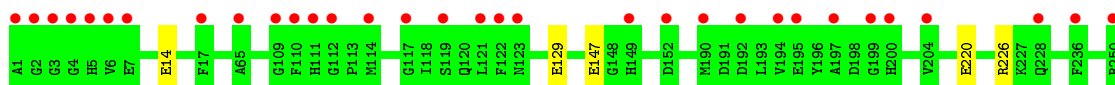
- Molecule 2: Cytochrome c1

Chain L: 

GLY  
HIS  
HIS  
HIS  
HIS  
HIS  
HIS

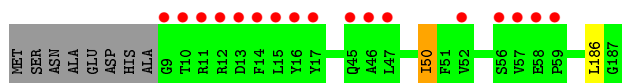
- Molecule 2: Cytochrome c1

Chain P: 

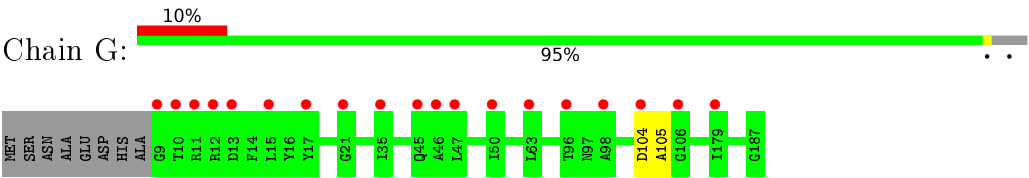


- Molecule 3: Ubiquinol-cytochrome c reductase iron-sulfur subunit

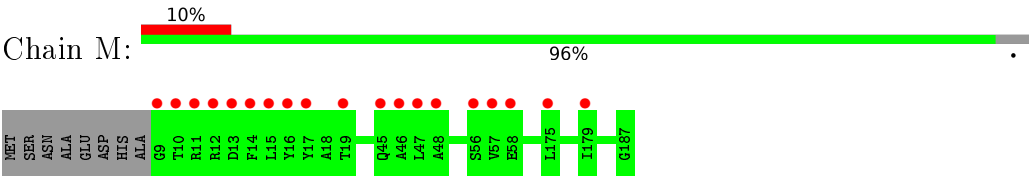
Chain C: 



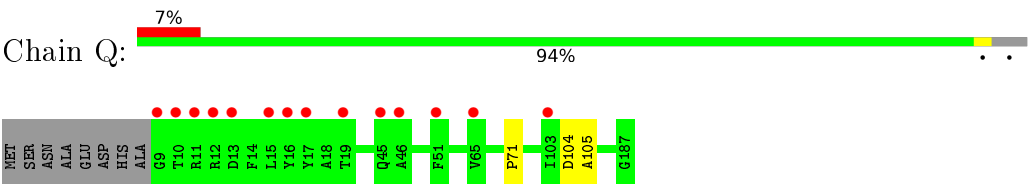
● Molecule 3: Ubiquinol-cytochrome c reductase iron-sulfur subunit



● Molecule 3: Ubiquinol-cytochrome c reductase iron-sulfur subunit



● Molecule 3: Ubiquinol-cytochrome c reductase iron-sulfur subunit



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.86Å 126.93Å 127.89Å 64.64° 87.69° 61.80°	Depositor
Resolution (Å)	37.74 – 3.00 37.72 – 3.00	Depositor EDS
% Data completeness (in resolution range)	96.2 (37.74-3.00) 66.8 (37.72-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.84 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.11rc1_2513: ???)	Depositor
R, $R_{free}$	0.211 , 0.243 0.203 , 0.243	Depositor DCC
$R_{free}$ test set	1820 reflections (1.80%)	DCC
Wilson B-factor (Å <sup>2</sup> )	58.3	Xtriage
Anisotropy	0.217	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 39.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.003 for h,h-k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	55093	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.91 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.4145e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SR, ANJ, LOP, FES, HEC, HEM, SMA

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	0.38	1/3565 (0.0%)	0.58	0/4891
1	E	0.38	1/3565 (0.0%)	0.58	0/4891
1	K	0.39	1/3565 (0.0%)	0.61	0/4891
1	O	0.37	1/3565 (0.0%)	0.58	0/4891
2	B	0.35	0/2010	0.57	0/2733
2	F	0.34	0/2010	0.58	1/2733 (0.0%)
2	L	0.35	0/2010	0.59	0/2733
2	P	0.33	0/2010	0.55	0/2733
3	C	0.35	0/1371	0.59	1/1868 (0.1%)
3	G	0.35	0/1371	0.57	0/1868
3	M	0.35	0/1371	0.58	0/1868
3	Q	0.36	0/1371	0.57	0/1868
All	All	0.37	4/27784 (0.0%)	0.58	2/37968 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	70	THR	C-N	7.44	1.48	1.34
1	K	70	THR	C-N	6.00	1.45	1.34
1	A	70	THR	C-N	5.91	1.45	1.34
1	O	70	THR	C-N	5.55	1.44	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	50	ILE	CG1-CB-CG2	5.21	122.87	111.40
2	F	158	ARG	NE-CZ-NH1	5.21	122.90	120.30

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	3435	3406	3420	3	0
1	E	3435	3406	3420	6	0
1	K	3435	3406	3420	6	0
1	O	3435	3406	3420	3	0
2	B	1953	1839	1848	1	0
2	F	1953	1839	1848	2	0
2	L	1953	1838	1848	3	0
2	P	1953	1839	1848	2	0
3	C	1341	1304	1307	1	0
3	G	1341	1304	1307	1	0
3	M	1341	1304	1307	0	0
3	Q	1341	1304	1307	2	0
4	A	86	60	60	2	0
4	E	86	60	60	2	0
4	K	86	60	60	2	0
4	O	86	60	60	2	0
5	A	37	42	42	0	0
5	E	37	42	42	0	0
5	K	37	42	42	0	0
5	O	37	42	42	0	0
6	A	39	39	39	0	0
6	E	39	39	39	3	0
6	K	39	39	40	1	0
6	O	39	39	39	1	0
7	A	1	0	0	0	0
7	B	1	0	0	0	0
7	F	1	0	0	0	0
7	K	1	0	0	0	0
7	L	1	0	0	0	0
7	P	1	0	0	0	0
8	A	45	67	67	0	0
8	E	45	67	67	0	0
8	K	45	67	67	1	0
8	O	45	67	67	0	0
9	B	43	32	30	2	0
9	F	43	32	30	2	0
9	L	43	32	30	3	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	P	43	32	30	2	0
10	C	4	0	0	0	0
10	G	4	0	0	0	0
10	M	4	0	0	0	0
10	Q	4	0	0	0	0
All	All	27938	27155	27253	42	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 1.

The worst 5 of 42 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:14:GLU:OE2	2:B:129:GLU:OE2	2.10	0.70
2:F:14:GLU:OE2	2:F:129:GLU:OE2	2.18	0.60
9:B:1001:HEC:HBC3	9:B:1001:HEC:HMC1	1.83	0.60
2:F:220:GLU:OE1	2:F:226:ARG:NH1	2.35	0.59
9:F:1001:HEC:HMC1	9:F:1001:HEC:HBC3	1.86	0.58

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	426/445 (96%)	421 (99%)	5 (1%)	0	100	100
1	E	426/445 (96%)	421 (99%)	5 (1%)	0	100	100
1	K	426/445 (96%)	421 (99%)	5 (1%)	0	100	100
1	O	426/445 (96%)	422 (99%)	4 (1%)	0	100	100
2	B	254/272 (93%)	246 (97%)	7 (3%)	1 (0%)	39	80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	254/272 (93%)	246 (97%)	7 (3%)	1 (0%)	39	80
2	L	254/272 (93%)	246 (97%)	7 (3%)	1 (0%)	39	80
2	P	254/272 (93%)	246 (97%)	7 (3%)	1 (0%)	39	80
3	C	177/187 (95%)	171 (97%)	6 (3%)	0	100	100
3	G	177/187 (95%)	171 (97%)	6 (3%)	0	100	100
3	M	177/187 (95%)	170 (96%)	7 (4%)	0	100	100
3	Q	177/187 (95%)	170 (96%)	7 (4%)	0	100	100
All	All	3428/3616 (95%)	3351 (98%)	73 (2%)	4 (0%)	56	90

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	147	GLU
2	B	147	GLU
2	F	147	GLU
2	P	147	GLU

### 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	353/366 (96%)	350 (99%)	3 (1%)	86	96
1	E	353/366 (96%)	350 (99%)	3 (1%)	86	96
1	K	353/366 (96%)	350 (99%)	3 (1%)	86	96
1	O	353/366 (96%)	350 (99%)	3 (1%)	86	96
2	B	203/216 (94%)	203 (100%)	0	100	100
2	F	203/216 (94%)	203 (100%)	0	100	100
2	L	203/216 (94%)	203 (100%)	0	100	100
2	P	203/216 (94%)	203 (100%)	0	100	100
3	C	138/144 (96%)	138 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	G	138/144 (96%)	138 (100%)	0	100	100
3	M	138/144 (96%)	138 (100%)	0	100	100
3	Q	138/144 (96%)	138 (100%)	0	100	100
All	All	2776/2904 (96%)	2764 (100%)	12 (0%)	93	98

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	199	TYR
1	K	104	PHE
1	O	104	PHE
1	E	108	VAL
1	K	199	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 34 ligands modelled in this entry, 6 are monoatomic - leaving 28 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	$\# Z  > 2$
4	HEM	A	1001	1	24,50,50	1.99	6 (25%)	16,82,82	1.36	2 (12%)
4	HEM	A	1002	1	24,50,50	1.87	4 (16%)	16,82,82	1.60	3 (18%)
5	SMA	A	1003	-	35,38,38	1.45	3 (8%)	39,52,52	2.32	11 (28%)
6	ANJ	A	1004	-	40,40,40	0.85	0	31,54,54	1.49	5 (16%)
8	LOP	A	1006	-	43,44,44	0.91	1 (2%)	44,49,49	1.15	4 (9%)
9	HEC	B	1001	2	24,50,50	2.31	3 (12%)	19,82,82	2.83	6 (31%)
10	FES	C	1001	3	0,4,4	0.00	-	0,4,4	0.00	-
4	HEM	E	1001	1	24,50,50	1.83	5 (20%)	16,82,82	1.68	3 (18%)
4	HEM	E	1002	1	24,50,50	1.87	5 (20%)	16,82,82	2.05	5 (31%)
5	SMA	E	1003	-	35,38,38	1.43	3 (8%)	39,52,52	2.15	11 (28%)
6	ANJ	E	1004	-	40,40,40	0.91	2 (5%)	31,54,54	1.05	2 (6%)
8	LOP	E	1005	-	43,44,44	0.93	1 (2%)	44,49,49	1.23	4 (9%)
9	HEC	F	1001	2	24,50,50	2.26	4 (16%)	19,82,82	2.58	7 (36%)
10	FES	G	1001	3	0,4,4	0.00	-	0,4,4	0.00	-
4	HEM	K	1001	1	24,50,50	2.11	5 (20%)	16,82,82	1.56	3 (18%)
4	HEM	K	1002	1	24,50,50	1.89	5 (20%)	16,82,82	1.86	4 (25%)
5	SMA	K	1003	-	35,38,38	1.42	4 (11%)	39,52,52	2.25	10 (25%)
6	ANJ	K	1004	-	40,40,40	0.90	1 (2%)	31,54,54	1.15	5 (16%)
8	LOP	K	1006	-	43,44,44	0.93	2 (4%)	44,49,49	1.11	5 (11%)
9	HEC	L	1001	2	24,50,50	2.32	4 (16%)	19,82,82	2.41	6 (31%)
10	FES	M	1001	3	0,4,4	0.00	-	0,4,4	0.00	-
4	HEM	O	1001	1	24,50,50	1.95	4 (16%)	16,82,82	1.60	3 (18%)
4	HEM	O	1002	1	24,50,50	1.88	4 (16%)	16,82,82	1.92	4 (25%)
5	SMA	O	1003	-	35,38,38	1.44	4 (11%)	39,52,52	2.29	10 (25%)
6	ANJ	O	1004	-	40,40,40	0.87	0	31,54,54	1.28	4 (12%)
8	LOP	O	1005	-	43,44,44	0.91	1 (2%)	44,49,49	1.20	3 (6%)
9	HEC	P	1001	2	24,50,50	2.31	5 (20%)	19,82,82	2.68	6 (31%)
10	FES	Q	1001	3	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
4	HEM	A	1001	1	-	0/6/54/54	0/0/8/8
4	HEM	A	1002	1	-	0/6/54/54	0/0/8/8
5	SMA	A	1003	-	-	0/33/34/34	0/2/2/2
6	ANJ	A	1004	-	-	1/39/55/55	0/1/2/2
8	LOP	A	1006	-	-	0/48/48/48	0/0/0/0
9	HEC	B	1001	2	-	0/6/54/54	0/0/8/8
10	FES	C	1001	3	-	0/0/4/4	0/1/1/1
4	HEM	E	1001	1	-	0/6/54/54	0/0/8/8
4	HEM	E	1002	1	-	0/6/54/54	0/0/8/8
5	SMA	E	1003	-	-	0/33/34/34	0/2/2/2
6	ANJ	E	1004	-	-	0/39/55/55	0/1/2/2
8	LOP	E	1005	-	-	0/48/48/48	0/0/0/0
9	HEC	F	1001	2	-	0/6/54/54	0/0/8/8
10	FES	G	1001	3	-	0/0/4/4	0/1/1/1
4	HEM	K	1001	1	-	0/6/54/54	0/0/8/8
4	HEM	K	1002	1	-	0/6/54/54	0/0/8/8
5	SMA	K	1003	-	-	0/33/34/34	0/2/2/2
6	ANJ	K	1004	-	-	1/39/55/55	0/1/2/2
8	LOP	K	1006	-	-	0/48/48/48	0/0/0/0
9	HEC	L	1001	2	-	0/6/54/54	0/0/8/8
10	FES	M	1001	3	-	0/0/4/4	0/1/1/1
4	HEM	O	1001	1	-	0/6/54/54	0/0/8/8
4	HEM	O	1002	1	-	0/6/54/54	0/0/8/8
5	SMA	O	1003	-	-	0/33/34/34	0/2/2/2
6	ANJ	O	1004	-	-	1/39/55/55	0/1/2/2
8	LOP	O	1005	-	-	0/48/48/48	0/0/0/0
9	HEC	P	1001	2	-	0/6/54/54	0/0/8/8
10	FES	Q	1001	3	-	0/0/4/4	0/1/1/1

The worst 5 of 76 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	B	1001	HEC	C3B-C2B	-6.79	1.33	1.40
9	P	1001	HEC	C3B-C2B	-6.58	1.33	1.40
9	L	1001	HEC	C3B-C2B	-6.01	1.34	1.40
9	F	1001	HEC	C3B-C2B	-5.99	1.34	1.40
9	L	1001	HEC	C3C-C2C	-5.84	1.34	1.40

The worst 5 of 126 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	B	1001	HEC	CBB-CAB-C3B	-7.29	111.40	127.34
9	P	1001	HEC	CBB-CAB-C3B	-7.18	111.64	127.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	F	1001	HEC	CBB-CAB-C3B	-6.87	112.32	127.34
9	B	1001	HEC	CBC-CAC-C3C	-6.29	113.58	127.34
9	L	1001	HEC	CBB-CAB-C3B	-6.22	113.73	127.34

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	O	1004	ANJ	C17-O6-C15-C16
6	K	1004	ANJ	C17-O6-C15-C16
6	A	1004	ANJ	C17-O6-C15-C16

There are no ring outliers.

13 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1001	HEM	1	0
4	A	1002	HEM	1	0
9	B	1001	HEC	2	0
4	E	1001	HEM	2	0
6	E	1004	ANJ	3	0
9	F	1001	HEC	2	0
4	K	1001	HEM	2	0
6	K	1004	ANJ	1	0
8	K	1006	LOP	1	0
9	L	1001	HEC	3	0
4	O	1001	HEM	2	0
6	O	1004	ANJ	1	0
9	P	1001	HEC	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(Å²)	Q<0.9
1	A	428/445 (96%)	0.43	28 (6%)	22	8	53, 88, 132, 178	0
1	E	428/445 (96%)	0.29	29 (6%)	20	7	49, 73, 119, 169	0
1	K	428/445 (96%)	0.17	15 (3%)	48	21	40, 63, 99, 154	0
1	O	428/445 (96%)	0.22	16 (3%)	45	19	49, 76, 119, 149	0
2	B	256/272 (94%)	0.43	24 (9%)	11	4	56, 85, 134, 199	0
2	F	256/272 (94%)	0.44	26 (10%)	9	3	57, 89, 130, 164	0
2	L	256/272 (94%)	0.42	23 (8%)	12	4	45, 75, 123, 167	0
2	P	256/272 (94%)	0.53	36 (14%)	4	1	64, 95, 130, 168	0
3	C	179/187 (95%)	0.68	17 (9%)	10	4	53, 87, 141, 184	0
3	G	179/187 (95%)	0.61	19 (10%)	8	3	67, 90, 133, 182	0
3	M	179/187 (95%)	0.71	19 (10%)	8	3	61, 85, 132, 172	0
3	Q	179/187 (95%)	0.49	14 (7%)	16	6	49, 86, 138, 190	0
All	All	3452/3616 (95%)	0.40	266 (7%)	16	6	40, 81, 128, 199	0

The worst 5 of 266 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	2	GLY	13.8
3	C	10	THR	11.4
2	L	2	GLY	10.2
3	Q	46	ALA	9.7
3	C	9	GLY	9.4

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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
8	LOP	E	1005	45/45	0.87	0.36	4.61	75,95,109,113	0
8	LOP	O	1005	45/45	0.83	0.34	2.61	78,100,116,117	0
10	FES	Q	1001	4/4	0.99	0.26	1.96	58,63,64,66	0
8	LOP	K	1006	45/45	0.85	0.29	1.53	64,85,101,106	0
10	FES	M	1001	4/4	0.99	0.26	1.38	72,72,75,77	0
5	SMA	K	1003	37/37	0.93	0.26	1.33	47,64,83,86	0
10	FES	G	1001	4/4	0.99	0.28	1.30	71,74,74,76	0
8	LOP	A	1006	45/45	0.78	0.33	1.30	80,104,114,116	0
4	HEM	E	1001	43/43	0.97	0.24	0.92	56,69,85,89	0
10	FES	C	1001	4/4	0.99	0.23	0.82	59,62,64,64	0
4	HEM	K	1001	43/43	0.97	0.24	0.78	51,62,79,83	0
6	ANJ	O	1004	39/39	0.86	0.26	0.78	67,84,107,108	0
5	SMA	O	1003	37/37	0.92	0.23	0.65	53,68,90,93	0
4	HEM	A	1002	43/43	0.94	0.29	0.64	80,96,114,126	0
4	HEM	O	1001	43/43	0.97	0.23	0.63	48,64,79,81	0
4	HEM	A	1001	43/43	0.96	0.26	0.62	61,72,87,95	0
9	HEC	F	1001	43/43	0.95	0.23	0.57	47,61,82,85	0
5	SMA	E	1003	37/37	0.92	0.22	0.46	47,67,88,88	0
4	HEM	O	1002	43/43	0.96	0.24	0.46	75,80,97,98	0
4	HEM	K	1002	43/43	0.96	0.24	0.40	64,72,89,96	0
7	SR	A	1005	1/1	0.83	0.17	0.34	133,133,133,133	0
9	HEC	P	1001	43/43	0.96	0.21	0.24	47,62,75,83	0
4	HEM	E	1002	43/43	0.97	0.22	0.23	65,78,94,98	0
6	ANJ	A	1004	39/39	0.85	0.24	0.12	77,89,107,108	0
7	SR	K	1005	1/1	0.86	0.17	0.12	121,121,121,121	0
9	HEC	B	1001	43/43	0.97	0.20	0.11	41,56,69,76	0
6	ANJ	K	1004	39/39	0.91	0.20	-0.11	49,71,100,101	0
9	HEC	L	1001	43/43	0.97	0.19	-0.22	40,55,69,69	0
5	SMA	A	1003	37/37	0.93	0.23	-0.26	56,69,83,89	0
6	ANJ	E	1004	39/39	0.87	0.20	-0.29	58,74,101,101	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
7	SR	F	1002	1/1	0.91	0.07	-1.35	132,132,132,132	0
7	SR	P	1002	1/1	0.95	0.05	-1.53	144,144,144,144	0
7	SR	L	1002	1/1	0.97	0.05	-1.65	113,113,113,113	0
7	SR	B	1002	1/1	0.95	0.05	-1.69	117,117,117,117	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.