



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 31, 2017 – 08:26 AM EDT

PDB ID : 3S8G  
Title : 1.8 Å structure of ba3 cytochrome c oxidase mutant (A120F) from *Thermus thermophilus* in lipid environment  
Authors : Tiefenbrunn, T.; Liu, W.; Chen, Y.; Katritch, V.; Stout, C.D.; Fee, J.A.; Cherezov, V.  
Deposited on : unknown  
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

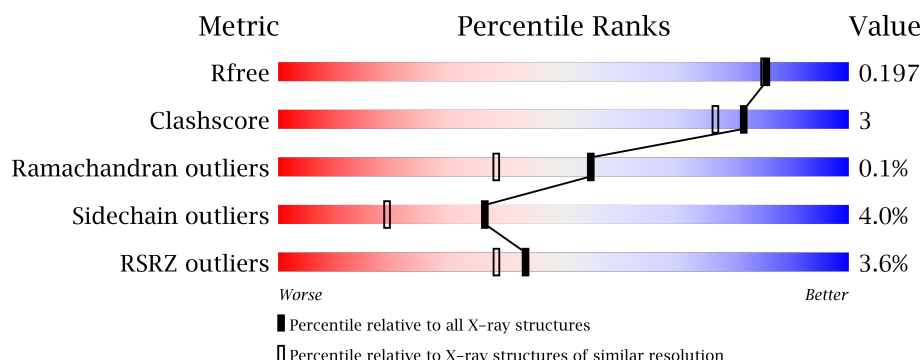
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	4827 (1.80-1.80)
Clashscore	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)
RSRZ outliers	101464	4906 (1.80-1.80)

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	569	<div> <div>4%</div> <div> <div></div> <div>90%</div> <div>7%</div> <div>..</div> </div> </div>
2	B	168	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>7%</div> <div>..</div> </div> </div>
3	C	34	<div> <div></div> <div> <div>76%</div> <div>9%</div> <div>6%</div> <div>9%</div> </div> </div>

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	OLC	A	564	-	-	-	X
8	OLC	A	565	-	-	-	X
8	OLC	A	566	-	-	-	X
8	OLC	A	567	-	-	-	X
8	OLC	A	569	-	-	-	X
8	OLC	A	570	-	-	-	X
8	OLC	A	572	-	-	-	X
8	OLC	A	573	-	-	-	X
8	OLC	A	574	-	-	-	X
8	OLC	A	575	-	-	-	X
8	OLC	A	576	-	-	-	X
8	OLC	A	577	-	-	-	X
8	OLC	B	169	-	-	-	X
8	OLC	B	170	-	-	-	X
8	OLC	B	171	-	-	-	X
8	OLC	B	172	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 6695 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Cytochrome c oxidase subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	554	Total	C	N	O	S	0	7	0
			4399	2988	700	694	17			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	EXPRESSION TAG	UNP Q5SJ79
A	-5	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	-4	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	-3	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	-2	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	-1	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	0	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	1	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	120	PHE	ALA	CONFLICT	UNP Q5SJ79

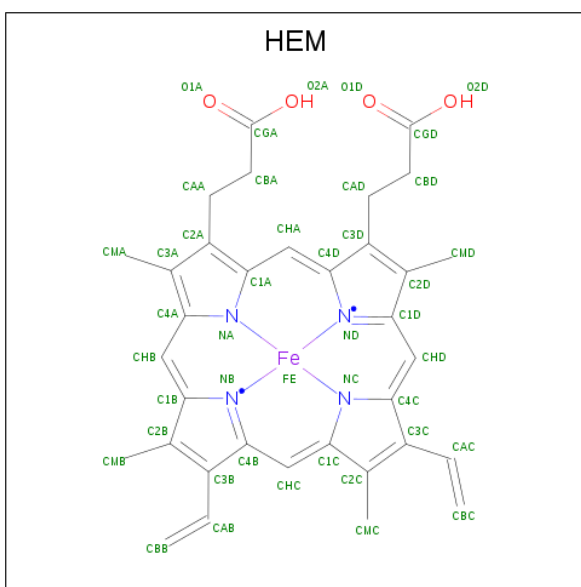
- Molecule 2 is a protein called Cytochrome c oxidase subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	166	Total	C	N	O	S	0	5	0
			1307	846	219	238	4			

- Molecule 3 is a protein called Cytochrome c oxidase polypeptide 2A.

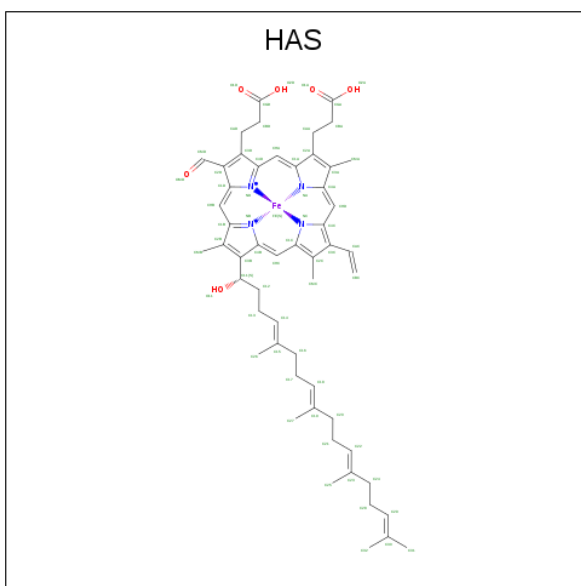
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	31	Total	C	N	O	0	0	0
			241	169	37	35			

- Molecule 4 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>4</sub>).



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

- Molecule 5 is HEME-AS (three-letter code: HAS) (formula:  $\text{C}_{54}\text{H}_{64}\text{FeN}_4\text{O}_6$ ).

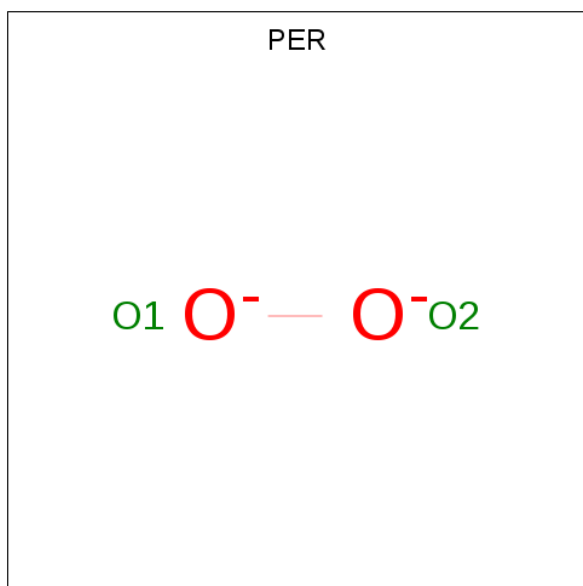


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	Fe	N	O	0	0
			65	54	1	4	6		

- Molecule 6 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

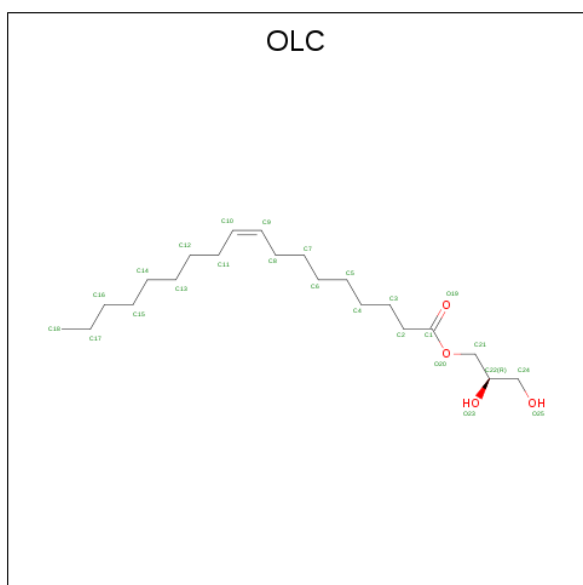
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cu 1 1	0	0

- Molecule 7 is PEROXIDE ION (three-letter code: PER) (formula: O<sub>2</sub>).



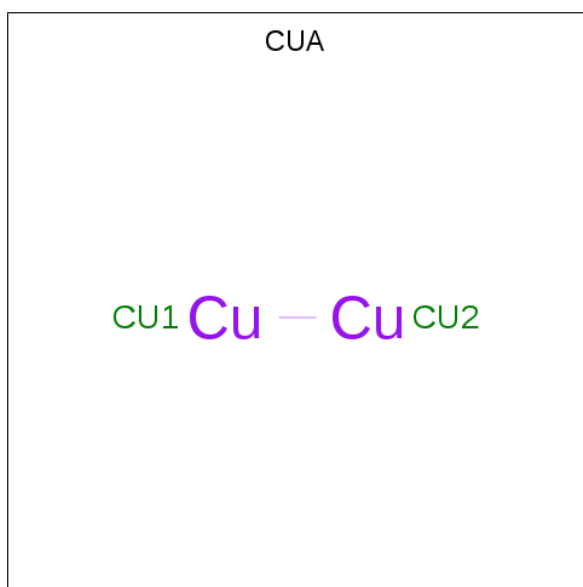
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total O 2 2	0	0

- Molecule 8 is (2R)-2,3-dihydroxypropyl (9Z)-octadec-9-enoate (three-letter code: OLC) (formula: C<sub>21</sub>H<sub>40</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			25	21	4		
8	A	1	Total	C	O	0	0
			25	21	4		
8	A	1	Total	C	O	0	0
			25	21	4		
8	A	1	Total	C	O	0	0
			23	19	4		
8	A	1	Total	C	O	0	0
			25	21	4		
8	A	1	Total	C	O	0	0
			18	14	4		
8	A	1	Total	C	O	0	0
			17	13	4		
8	A	1	Total	C	O	0	0
			8	4	4		
8	A	1	Total	C	O	0	0
			15	11	4		
8	A	1	Total	C	O	0	0
			20	16	4		
8	A	1	Total	C	O	0	0
			25	21	4		
8	A	1	Total	C	O	0	0
			21	17	4		
8	A	1	Total	C	O	0	0
			16	12	4		
8	A	1	Total	C	O	0	0
			25	21	4		
8	A	1	Total	C	O	0	0
			12	8	4		
8	B	1	Total	C	O	0	0
			25	21	4		
8	B	1	Total	C	O	0	0
			25	21	4		
8	B	1	Total	C	O	0	0
			25	21	4		
8	B	1	Total	C	O	0	0
			21	19	2		
8	C	1	Total	C	O	0	0
			14	10	4		

- Molecule 9 is DINUCLEAR COPPER ION (three-letter code: CUA) (formula: Cu<sub>2</sub>).



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

- Molecule 10 is water.

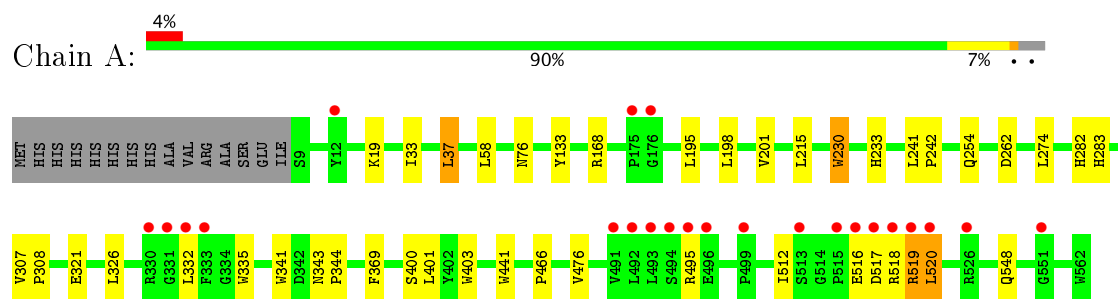
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	137	Total O 137 137	0	0
10	B	81	Total O 81 81	0	0
10	C	7	Total O 7 7	0	0



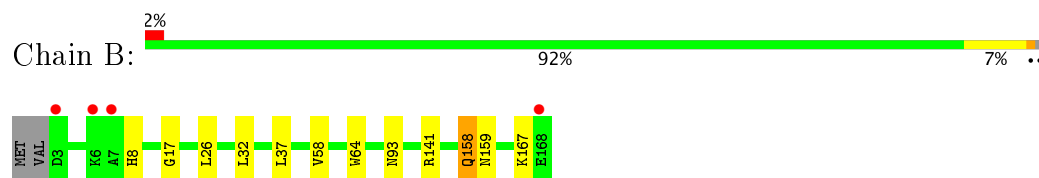
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

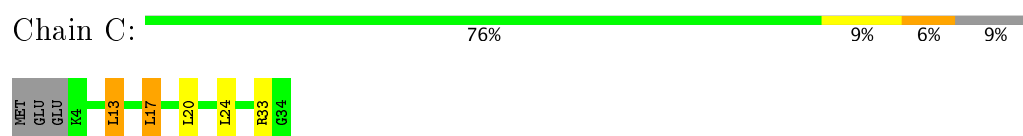
- Molecule 1: Cytochrome c oxidase subunit 1



- Molecule 2: Cytochrome c oxidase subunit 2



- Molecule 3: Cytochrome c oxidase polypeptide 2A



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	144.96Å 98.64Å 95.06Å 90.00° 128.07° 90.00°	Depositor
Resolution (Å)	39.48 – 1.80 39.48 – 1.80	Depositor EDS
% Data completeness (in resolution range)	97.8 (39.48-1.80) 97.8 (39.48-1.80)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.16 (at 1.79Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.175 , 0.196 0.175 , 0.197	Depositor DCC
$R_{free}$ test set	4758 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	21.3	Xtriage
Anisotropy	0.131	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 66.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.017 for -h+k-l,-l,-k 0.002 for -h-k-l,l,k 0.018 for -h-2*l,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6695	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: OLC, PER, CUA, HEM, HAS, CU

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	0.68	0/4584	0.64	1/6292 (0.0%)
2	B	0.69	0/1364	0.61	0/1862
3	C	0.73	0/247	0.72	0/335
All	All	0.69	0/6195	0.64	1/8489 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	233	HIS	N-CA-CB	5.24	120.04	110.60

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	4399	0	4496	21	0
2	B	1307	0	1277	9	0
3	C	241	0	267	4	0
4	A	43	0	30	3	0
5	A	65	0	62	1	0
6	A	1	0	0	0	0
7	A	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	A	300	0	434	8	0
8	B	96	0	153	7	0
8	C	14	0	17	0	0
9	B	2	0	0	0	0
10	A	137	0	0	2	0
10	B	81	0	0	1	1
10	C	7	0	0	2	0
All	All	6695	0	6736	39	1

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 3.

All (39) 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:141:ARG:HH12	8:B:170:OLC:H24A	1.19	1.05
1:A:168:ARG:HH22	8:A:572:OLC:H6A	1.26	0.97
2:B:158[B]:GLN:HG2	2:B:159[B]:ASN:OD1	1.80	0.80
3:C:33:ARG:HD2	10:C:36:HOH:O	1.89	0.72
1:A:168:ARG:NH2	8:A:572:OLC:H6A	2.02	0.72
8:B:170:OLC:H6	3:C:33:ARG:HE	1.55	0.71
2:B:141:ARG:NH1	8:B:170:OLC:H24A	2.02	0.68
1:A:198:LEU:HA	1:A:201[B]:VAL:HG22	1.75	0.66
2:B:93:ASN:O	10:B:1083:HOH:O	2.15	0.64
8:A:578:OLC:H21A	8:B:170:OLC:H14	1.83	0.60
3:C:33:ARG:CD	10:C:36:HOH:O	2.48	0.59
4:A:800:HEM:HBC2	4:A:800:HEM:CMC	2.36	0.56
4:A:800:HEM:HBC2	4:A:800:HEM:HMC2	1.90	0.53
1:A:517:ASP:CG	1:A:520:LEU:HB2	2.31	0.51
1:A:254[B]:GLN:OE1	10:A:1077:HOH:O	2.19	0.51
2:B:58:VAL:HG22	2:B:64:TRP:HB2	1.92	0.51
1:A:282:HIS:CD2	1:A:283:HIS:CD2	3.00	0.50
1:A:400:SER:HA	1:A:403:TRP:NE1	2.27	0.49
1:A:517:ASP:HA	1:A:519:ARG:HE	1.78	0.49
5:A:801:HAS:HBC1	5:A:801:HAS:HMC1	1.94	0.48
1:A:476:VAL:HG22	8:A:566:OLC:H16A	1.96	0.47
2:B:32:LEU:HD21	8:B:169:OLC:H7A	1.97	0.45
3:C:13:LEU:HD22	3:C:17:LEU:HD22	1.98	0.45
1:A:321:GLU:HA	1:A:335:TRP:CE3	2.53	0.44
1:A:33:ILE:O	1:A:37:LEU:HD13	2.18	0.43
8:A:575:OLC:H8	8:A:575:OLC:H11	1.85	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:19:LYS:HZ1	8:A:576:OLC:H21A	1.84	0.43
1:A:76:ASN:HB3	4:A:800:HEM:CAC	2.49	0.43
1:A:441:TRP:HB3	1:A:466:PRO:HB3	2.01	0.43
2:B:141:ARG:HH12	8:B:170:OLC:C24	2.07	0.43
2:B:17:GLY:HA3	8:B:171:OLC:H3	2.01	0.42
1:A:307:VAL:N	1:A:308:PRO:HD2	2.34	0.42
1:A:548:GLN:NE2	10:A:1201:HOH:O	2.52	0.42
1:A:230:TRP:CD1	1:A:230:TRP:C	2.93	0.41
1:A:241:LEU:N	1:A:242:PRO:CD	2.83	0.41
8:A:574:OLC:H8A	8:A:574:OLC:H13A	2.03	0.41
1:A:512:ILE:HG21	2:B:8:HIS:HB2	2.01	0.41
1:A:341:TRP:HB3	8:A:567:OLC:H24A	2.03	0.41
1:A:343:ASN:HA	1:A:344:PRO:HD2	1.99	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
10:B:1062:HOH:O	10:B:1063:HOH:O[2_556]	1.54	0.66

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	558/569 (98%)	542 (97%)	15 (3%)	1 (0%)	51	35
2	B	169/168 (101%)	167 (99%)	2 (1%)	0	100	100
3	C	29/34 (85%)	29 (100%)	0	0	100	100
All	All	756/771 (98%)	738 (98%)	17 (2%)	1 (0%)	55	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	516	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	453/464 (98%)	437 (96%)	16 (4%)	41	24
2	B	138/138 (100%)	133 (96%)	5 (4%)	40	23
3	C	24/27 (89%)	20 (83%)	4 (17%)	2	0
All	All	615/629 (98%)	590 (96%)	25 (4%)	36	18

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	LEU
1	A	58	LEU
1	A	133	TYR
1	A	195	LEU
1	A	215	LEU
1	A	230	TRP
1	A	262	ASP
1	A	274	LEU
1	A	326	LEU
1	A	332	LEU
1	A	369	PHE
1	A	401	LEU
1	A	495	ARG
1	A	518	ARG
1	A	519	ARG
1	A	520	LEU
2	B	26	LEU
2	B	37	LEU
2	B	158[A]	GLN
2	B	158[B]	GLN
2	B	167	LYS
3	C	13	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	17	LEU
3	C	20	LEU
3	C	24	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	407	ASN
2	B	8	HIS
2	B	40	HIS

### 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 25 ligands modelled in this entry, 1 is monoatomic - leaving 24 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	$\# Z  > 2$
7	PER	A	563	5,6	0,1,1	0.00	-	0,0,0	0.00	-
8	OLC	A	564	-	24,24,24	0.44	0	25,25,25	0.69	0
8	OLC	A	565	-	24,24,24	0.44	0	25,25,25	0.57	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	OLC	A	566	-	24,24,24	0.44	0	25,25,25	0.62	0
8	OLC	A	567	-	22,22,24	0.46	0	23,23,25	0.77	1 (4%)
8	OLC	A	568	-	24,24,24	0.45	0	25,25,25	0.57	0
8	OLC	A	569	-	17,17,24	0.56	0	18,18,25	0.66	0
8	OLC	A	570	-	16,16,24	0.51	0	17,17,25	0.61	0
8	OLC	A	571	-	7,7,24	0.63	0	6,7,25	0.42	0
8	OLC	A	572	-	14,14,24	0.58	0	15,15,25	0.55	0
8	OLC	A	573	-	19,19,24	0.50	0	20,20,25	0.75	0
8	OLC	A	574	-	24,24,24	0.41	0	25,25,25	0.69	0
8	OLC	A	575	-	20,20,24	0.47	0	21,21,25	0.73	0
8	OLC	A	576	-	15,15,24	0.54	0	16,16,25	0.63	0
8	OLC	A	577	-	24,24,24	0.42	0	25,25,25	0.64	0
8	OLC	A	578	-	11,11,24	0.61	0	12,12,25	0.59	0
4	HEM	A	800	1	28,50,50	2.53	9 (32%)	17,82,82	1.37	3 (17%)
5	HAS	A	801	1,7	55,72,72	4.36	16 (29%)	48,109,109	2.40	11 (22%)
8	OLC	B	169	-	24,24,24	0.45	0	25,25,25	0.63	0
8	OLC	B	170	-	24,24,24	0.52	0	25,25,25	0.64	1 (4%)
8	OLC	B	171	-	24,24,24	0.46	0	25,25,25	0.55	0
8	OLC	B	172	-	20,20,24	0.47	0	20,20,25	0.60	0
9	CUA	B	802	2	0,1,1	0.00	-	0,0,0	0.00	-
8	OLC	C	35	-	13,13,24	0.59	0	14,14,25	0.68	0

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
7	PER	A	563	5,6	-	0/0/0/0	0/0/0/0
8	OLC	A	564	-	-	0/24/24/24	0/0/0/0
8	OLC	A	565	-	-	0/24/24/24	0/0/0/0
8	OLC	A	566	-	-	0/24/24/24	0/0/0/0
8	OLC	A	567	-	-	0/22/22/24	0/0/0/0
8	OLC	A	568	-	-	0/24/24/24	0/0/0/0
8	OLC	A	569	-	-	0/17/17/24	0/0/0/0
8	OLC	A	570	-	-	0/16/16/24	0/0/0/0
8	OLC	A	571	-	-	0/6/6/24	0/0/0/0
8	OLC	A	572	-	-	0/14/14/24	0/0/0/0
8	OLC	A	573	-	-	0/19/19/24	0/0/0/0
8	OLC	A	574	-	-	0/24/24/24	0/0/0/0
8	OLC	A	575	-	-	0/20/20/24	0/0/0/0
8	OLC	A	576	-	-	0/15/15/24	0/0/0/0

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	OLC	A	577	-	-	0/24/24/24	0/0/0/0
8	OLC	A	578	-	-	0/11/11/24	0/0/0/0
4	HEM	A	800	1	-	0/6/54/54	0/0/8/8
5	HAS	A	801	1,7	-	0/35/122/122	0/0/8/8
8	OLC	B	169	-	-	0/24/24/24	0/0/0/0
8	OLC	B	170	-	-	0/24/24/24	0/0/0/0
8	OLC	B	171	-	-	0/24/24/24	0/0/0/0
8	OLC	B	172	-	-	0/19/19/24	0/0/0/0
9	CUA	B	802	2	-	0/0/0/0	0/0/0/0
8	OLC	C	35	-	-	0/13/13/24	0/0/0/0

All (25) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	801	HAS	CHD-C4A	-19.55	1.32	1.51
5	A	801	HAS	C1B-NB	-10.62	1.35	1.49
5	A	801	HAS	C4B-NB	-10.49	1.35	1.49
5	A	801	HAS	CHC-C4B	-9.50	1.35	1.53
5	A	801	HAS	CHA-C1A	-7.39	1.33	1.51
4	A	800	HEM	C3B-CAB	-6.76	1.34	1.47
5	A	801	HAS	C3C-CAC	-6.12	1.35	1.47
5	A	801	HAS	CHB-C1B	-4.81	1.33	1.53
5	A	801	HAS	CHD-C4C	-3.67	1.47	1.51
5	A	801	HAS	C1D-ND	-3.45	1.45	1.49
5	A	801	HAS	C4D-ND	-3.20	1.45	1.49
4	A	800	HEM	C1D-CHD	-2.65	1.33	1.40
5	A	801	HAS	CHA-C4D	-2.51	1.48	1.53
4	A	800	HEM	C4A-CHB	-2.51	1.33	1.40
4	A	800	HEM	C4B-CHC	-2.37	1.33	1.40
5	A	801	HAS	C1C-C2C	-2.32	1.34	1.38
5	A	801	HAS	CBC-CAC	2.99	1.49	1.28
4	A	800	HEM	CBB-CAB	3.18	1.51	1.28
4	A	800	HEM	C1A-CHA	3.90	1.50	1.40
4	A	800	HEM	C1A-NA	4.11	1.45	1.36
4	A	800	HEM	C3C-C2C	4.88	1.46	1.40
4	A	800	HEM	C3B-C2B	6.03	1.48	1.40
5	A	801	HAS	C3C-C2C	6.71	1.49	1.40
5	A	801	HAS	C4A-C3A	7.10	1.47	1.38
5	A	801	HAS	C1A-C2A	7.90	1.49	1.38

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	801	HAS	CHD-C4A-C3A	-5.65	120.09	129.53
5	A	801	HAS	CAA-C2A-C1A	-4.32	124.26	127.30
5	A	801	HAS	CHD-C4C-C3C	-3.31	125.27	129.61
5	A	801	HAS	CHC-C1C-C2C	-3.21	123.90	129.45
5	A	801	HAS	C4C-C3C-C2C	-2.98	99.68	104.13
5	A	801	HAS	CAA-CBA-CGA	-2.98	107.57	112.66
5	A	801	HAS	CBD-CAD-C3D	-2.84	109.10	114.28
4	A	800	HEM	CMA-C3A-C4A	-2.45	124.69	128.46
4	A	800	HEM	C4C-C3C-C2C	-2.45	105.19	106.90
4	A	800	HEM	C3B-C4B-NB	-2.14	106.44	109.21
8	B	170	OLC	O20-C21-C22	2.16	116.30	105.72
8	A	567	OLC	C21-O20-C1	2.21	123.77	117.13
5	A	801	HAS	C1B-CHB-C1D	3.29	126.18	116.30
5	A	801	HAS	CHA-C4D-ND	5.70	121.51	110.75
5	A	801	HAS	C4A-CHD-C4C	6.52	127.42	112.83
5	A	801	HAS	CHC-C4B-NB	8.36	126.52	110.75

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

12 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	566	OLC	1	0
8	A	567	OLC	1	0
8	A	572	OLC	2	0
8	A	574	OLC	1	0
8	A	575	OLC	1	0
8	A	576	OLC	1	0
8	A	578	OLC	1	0
4	A	800	HEM	3	0
5	A	801	HAS	1	0
8	B	169	OLC	1	0
8	B	170	OLC	5	0
8	B	171	OLC	1	0

## 5.7 Other polymers ⓘ

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<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	554/569 (97%)	-0.22	23 (4%) 37 31	11, 21, 43, 67	0
2	B	166/168 (98%)	-0.39	4 (2%) 59 55	12, 22, 38, 53	0
3	C	31/34 (91%)	-0.70	0 100 100	16, 21, 31, 41	0
All	All	751/771 (97%)	-0.28	27 (3%) 43 38	11, 21, 42, 67	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	495	ARG	5.3
1	A	515	PRO	5.0
1	A	333	PHE	4.9
1	A	496	GLU	4.4
1	A	516	GLU	4.2
1	A	513	SER	4.1
1	A	517	ASP	4.0
1	A	330[A]	ARG	4.0
1	A	176	GLY	3.6
1	A	332	LEU	3.6
2	B	6	LYS	3.5
1	A	493	LEU	3.5
1	A	331	GLY	3.4
1	A	491	VAL	3.0
1	A	492	LEU	2.9
1	A	519	ARG	2.9
2	B	168	GLU	2.9
1	A	12	TYR	2.7
1	A	520	LEU	2.6
1	A	494	SER	2.4
1	A	175	PRO	2.4
2	B	3	ASP	2.2
1	A	518	ARG	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	551	GLY	2.2
1	A	499	PRO	2.1
2	B	7	ALA	2.1
1	A	526	ARG	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	OLC	A	573	20/25	0.58	0.24	45.52	59,63,64,64	0
8	OLC	A	570	17/25	0.52	0.27	42.08	63,65,74,74	0
8	OLC	A	565	25/25	0.57	0.25	13.96	66,70,79,80	0
8	OLC	A	576	16/25	0.61	0.30	13.51	67,70,72,72	0
8	OLC	B	172	21/25	0.68	0.23	11.13	59,61,65,66	0
8	OLC	A	566	25/25	0.58	0.22	9.23	54,64,73,73	0
8	OLC	A	564	25/25	0.71	0.21	8.90	37,46,64,65	0
8	OLC	A	577	25/25	0.66	0.22	8.88	52,55,58,59	0
8	OLC	A	569	18/25	0.41	0.23	8.25	56,61,70,70	0
8	OLC	A	574	25/25	0.74	0.18	8.22	53,56,57,57	0
8	OLC	B	169	25/25	0.77	0.18	8.14	56,59,61,62	0
8	OLC	A	572	15/25	0.63	0.38	6.29	54,59,65,65	0
8	OLC	B	171	25/25	0.76	0.21	5.95	59,61,66,67	0
8	OLC	A	575	21/25	0.66	0.20	3.76	57,63,71,72	0
8	OLC	B	170	25/25	0.58	0.23	3.02	60,65,68,69	0
8	OLC	A	567	23/25	0.89	0.14	2.49	28,36,54,55	0
8	OLC	A	568	25/25	0.47	0.30	1.48	74,75,83,83	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	PER	A	563	2/2	0.97	0.13	0.80	14,14,14,15	0
5	HAS	A	801	65/65	0.99	0.11	0.32	8,13,26,30	0
4	HEM	A	800	43/43	0.99	0.08	-0.37	6,11,14,20	0
9	CUA	B	802	2/2	1.00	0.06	-2.46	14,14,14,14	0
8	OLC	A	571	8/25	0.84	0.25	-	59,61,63,64	0
6	CU	A	803	1/1	1.00	0.07	-	15,15,15,15	0
8	OLC	C	35	14/25	0.63	0.28	-	77,78,79,79	0
8	OLC	A	578	12/25	0.45	0.31	-	74,76,77,78	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.