



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:25 PM GMT

PDB ID : 3ZBY  
Title : Ligand-free structure of CYP142 from Mycobacterium smegmatis  
Authors : Garcia-Fernandez, E.; Frank, D.J.; Galan, B.; Kells, P.M.; Podust, L.M.; Garcia, J.L.; Ortiz de Montellano, P.R.  
Deposited on : 2012-11-13  
Resolution : 1.93 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

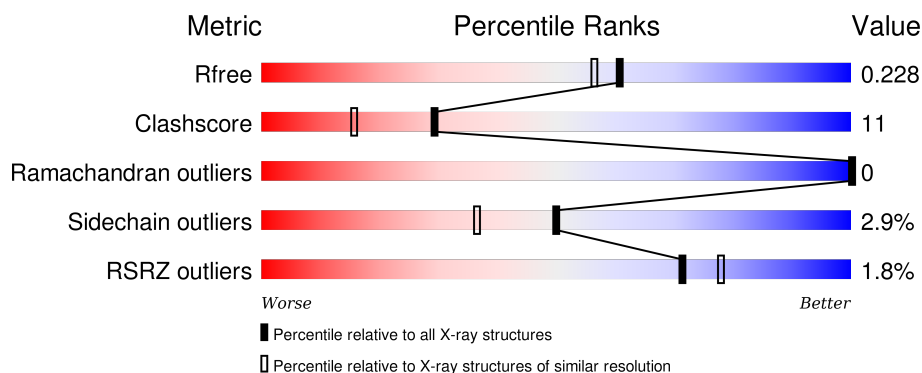
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.93 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	2910 (1.96-1.92)
Clashscore	102246	3095 (1.96-1.92)
Ramachandran outliers	100387	3062 (1.96-1.92)
Sidechain outliers	100360	3062 (1.96-1.92)
RSRZ outliers	91569	2915 (1.96-1.92)

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	407	 2% 84% 13% ..
1	B	407	 % 82% 14% ..
1	C	407	 2% 83% 14% ..
1	D	407	 % 82% 14% ..
1	E	407	 2% 83% 13% ..

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	407	

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
3	BCD	A	1403	X	-	-	X
3	BCD	B	1403	X	-	X	-
3	BCD	C	1403	X	-	X	-
3	BCD	D	1403	X	-	-	-
3	BCD	E	1403	X	-	X	-
3	BCD	F	1403	X	-	-	-
4	SO4	A	1404	-	-	X	X
4	SO4	B	1408	-	-	X	X
4	SO4	C	1406	-	-	X	X
4	SO4	D	1406	-	-	X	X
4	SO4	E	1406	-	-	X	X
4	SO4	F	1406	-	-	X	X
5	EDO	A	1406	-	-	-	X
5	EDO	B	1409	-	-	-	X
5	EDO	D	1407	-	-	-	X
5	EDO	E	1407	-	-	-	X
5	EDO	F	1407	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 22588 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 P 450 HEME-THIOLATE PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	402	Total	C	N	O	S	0	5	0
			3171	1990	561	595	25			
1	B	404	Total	C	N	O	S	0	5	0
			3181	1994	559	604	24			
1	C	402	Total	C	N	O	S	0	7	0
			3180	1995	561	600	24			
1	D	402	Total	C	N	O	S	0	7	0
			3205	2005	569	606	25			
1	E	402	Total	C	N	O	S	0	6	0
			3199	1999	574	602	24			
1	F	402	Total	C	N	O	S	0	6	0
			3186	1996	563	603	24			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	402	HIS	-	EXPRESSION TAG	UNP A0R4Q6
A	403	HIS	-	EXPRESSION TAG	UNP A0R4Q6
A	404	HIS	-	EXPRESSION TAG	UNP A0R4Q6
A	405	HIS	-	EXPRESSION TAG	UNP A0R4Q6
A	406	HIS	-	EXPRESSION TAG	UNP A0R4Q6
A	407	HIS	-	EXPRESSION TAG	UNP A0R4Q6
B	402	HIS	-	EXPRESSION TAG	UNP A0R4Q6
B	403	HIS	-	EXPRESSION TAG	UNP A0R4Q6
B	404	HIS	-	EXPRESSION TAG	UNP A0R4Q6
B	405	HIS	-	EXPRESSION TAG	UNP A0R4Q6
B	406	HIS	-	EXPRESSION TAG	UNP A0R4Q6
B	407	HIS	-	EXPRESSION TAG	UNP A0R4Q6
C	402	HIS	-	EXPRESSION TAG	UNP A0R4Q6
C	403	HIS	-	EXPRESSION TAG	UNP A0R4Q6
C	404	HIS	-	EXPRESSION TAG	UNP A0R4Q6
C	405	HIS	-	EXPRESSION TAG	UNP A0R4Q6
C	406	HIS	-	EXPRESSION TAG	UNP A0R4Q6

*Continued on next page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	407	HIS	-	EXPRESSION TAG	UNP A0R4Q6
D	402	HIS	-	EXPRESSION TAG	UNP A0R4Q6
D	403	HIS	-	EXPRESSION TAG	UNP A0R4Q6
D	404	HIS	-	EXPRESSION TAG	UNP A0R4Q6
D	405	HIS	-	EXPRESSION TAG	UNP A0R4Q6
D	406	HIS	-	EXPRESSION TAG	UNP A0R4Q6
D	407	HIS	-	EXPRESSION TAG	UNP A0R4Q6
E	402	HIS	-	EXPRESSION TAG	UNP A0R4Q6
E	403	HIS	-	EXPRESSION TAG	UNP A0R4Q6
E	404	HIS	-	EXPRESSION TAG	UNP A0R4Q6
E	405	HIS	-	EXPRESSION TAG	UNP A0R4Q6
E	406	HIS	-	EXPRESSION TAG	UNP A0R4Q6
E	407	HIS	-	EXPRESSION TAG	UNP A0R4Q6
F	402	HIS	-	EXPRESSION TAG	UNP A0R4Q6
F	403	HIS	-	EXPRESSION TAG	UNP A0R4Q6
F	404	HIS	-	EXPRESSION TAG	UNP A0R4Q6
F	405	HIS	-	EXPRESSION TAG	UNP A0R4Q6
F	406	HIS	-	EXPRESSION TAG	UNP A0R4Q6
F	407	HIS	-	EXPRESSION TAG	UNP A0R4Q6

- 
- Chemical structure of HEM (heme) showing a central iron atom coordinated by four nitrogen atoms in a porphyrin-like ring. The structure includes various side chains and a central heme group.

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

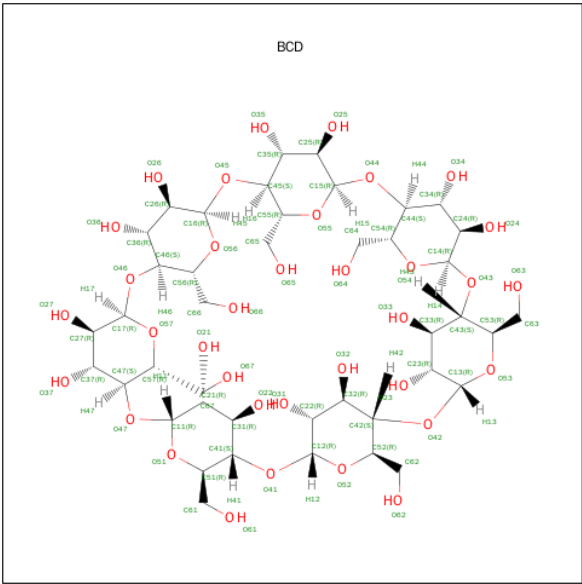


WORLD WIDE  
PDB  
PROTEIN DATA BANK

Continued from previous page...

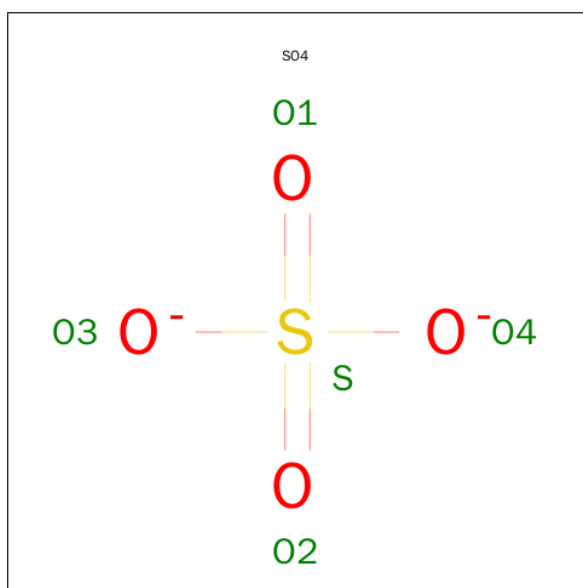
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	E	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	F	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is SUGAR (BETA-CYCLODEXTRIN) (three-letter code: BCD) (formula: C<sub>42</sub>H<sub>70</sub>O<sub>35</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			77	42	35		
3	B	1	Total	C	O	0	0
			77	42	35		
3	C	1	Total	C	O	0	0
			77	42	35		
3	D	1	Total	C	O	0	0
			77	42	35		
3	E	1	Total	C	O	0	0
			77	42	35		
3	F	1	Total	C	O	0	0
			77	42	35		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



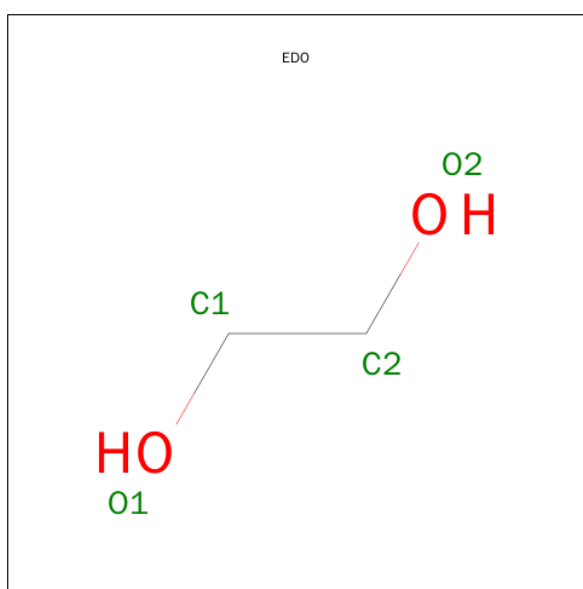
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	E	1	Total	O	S	0	0
			5	4	1		
4	E	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	E	1	Total	O	S	0	0
			5	4	1		
4	F	1	Total	O	S	0	0
			5	4	1		
4	F	1	Total	O	S	0	0
			5	4	1		
4	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	D	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	F	1	Total	C	O	0	0
			4	2	2		

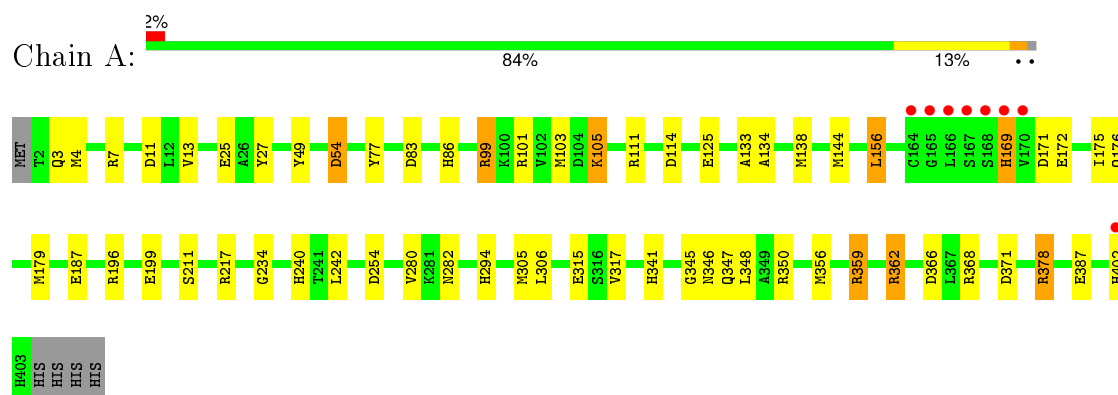
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	448	Total 448	O 448	0	0
6	B	439	Total 439	O 439	0	0
6	C	457	Total 457	O 457	0	0
6	D	431	Total 431	O 431	0	0
6	E	445	Total 445	O 445	0	0
6	F	417	Total 417	O 417	0	0

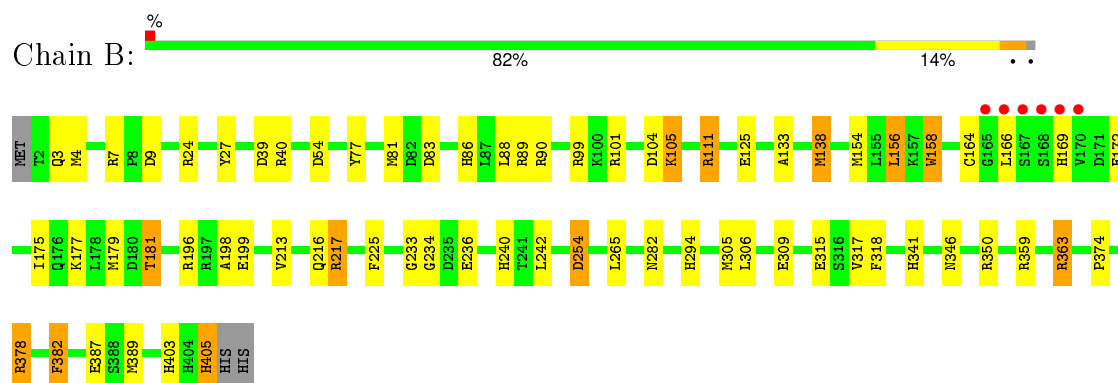
### 3 Residue-property plots [i](#)

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

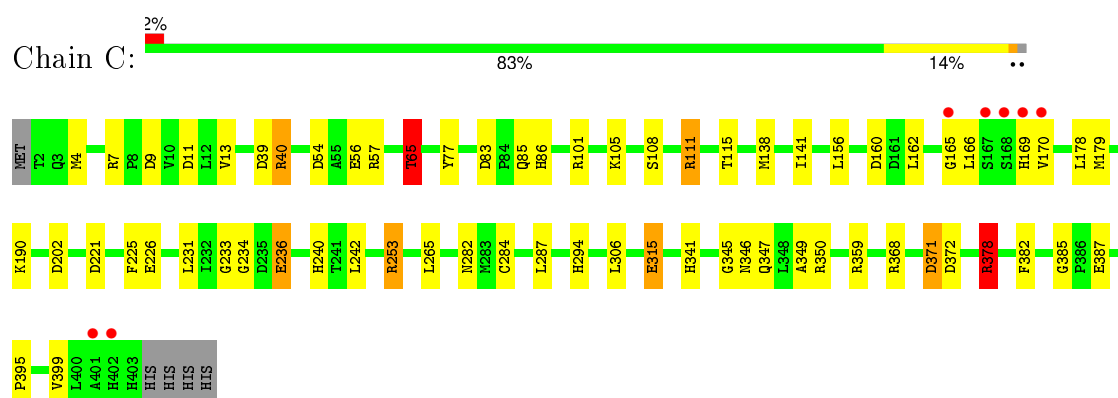
#### • Molecule 1: P450 HEME-THIOLATE PROTEIN



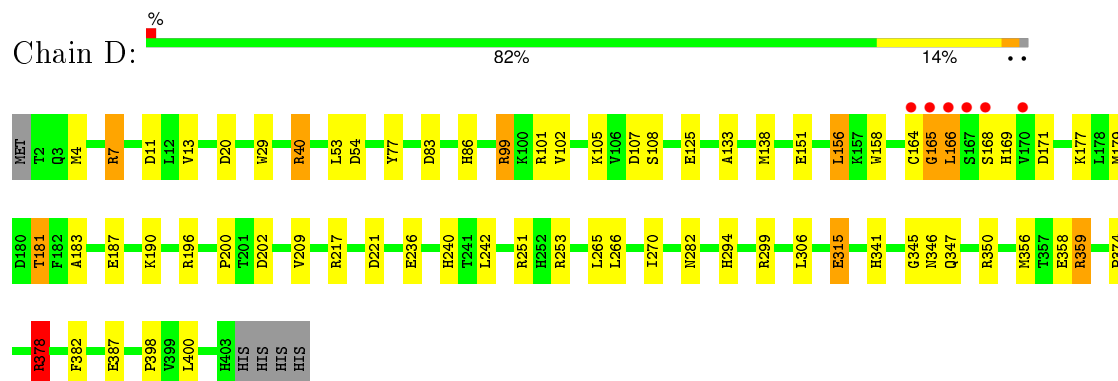
#### • Molecule 1: P450 HEME-THIOLATE PROTEIN



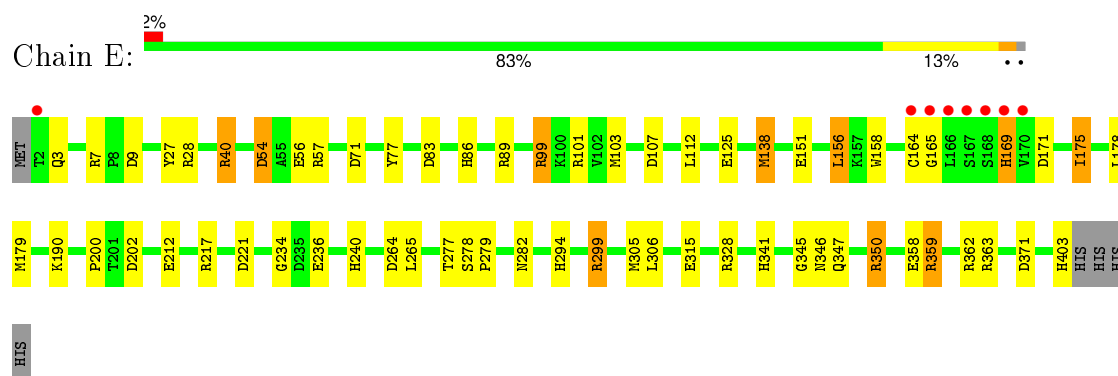
#### • Molecule 1: P450 HEME-THIOLATE PROTEIN



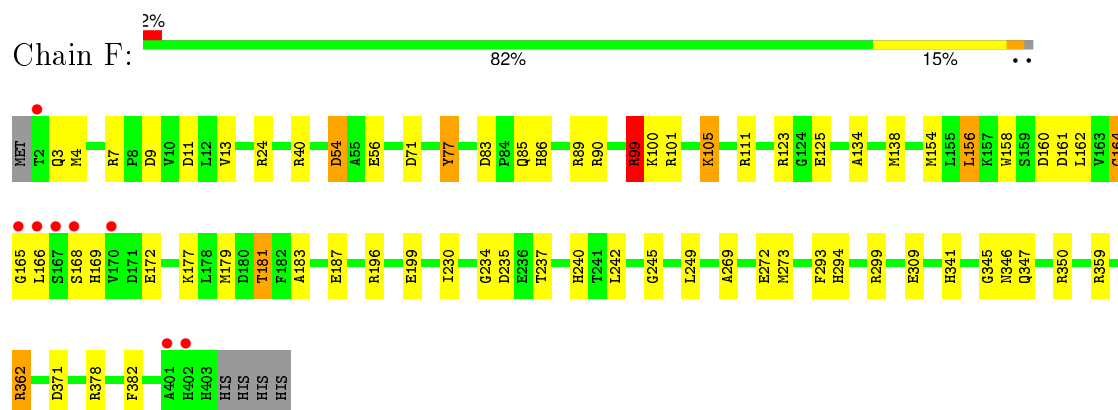
- Molecule 1: P450 HEME-THIOLATE PROTEIN



- Molecule 1: P450 HEME-THIOLATE PROTEIN



- Molecule 1: P450 HEME-THIOLATE PROTEIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.05Å 162.85Å 266.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	266.44 – 1.93 69.49 – 1.93	Depositor EDS
% Data completeness (in resolution range)	94.3 (266.44-1.93) 94.3 (69.49-1.93)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.16 (at 1.92Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.190 , 0.230 0.189 , 0.228	Depositor DCC
$R_{free}$ test set	14351 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 33.4	EDS
Estimated twinning fraction	0.460 for -1/2*h+1/2*k,3/2*h+1/2*k,-l 0.467 for -1/2*h-1/2*k,-3/2*h+1/2*k,-l 0.467 for 1/2*h+1/2*k,3/2*h-1/2*k,-l 0.467 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.467 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	3 of 283950 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	22588	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: BCD, HEM, SO4, EDO

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	1.17	4/3237 (0.1%)	1.01	9/4391 (0.2%)
1	B	1.18	7/3251 (0.2%)	1.06	16/4414 (0.4%)
1	C	1.20	7/3255 (0.2%)	1.06	14/4416 (0.3%)
1	D	1.17	5/3274 (0.2%)	1.02	12/4439 (0.3%)
1	E	1.20	6/3265 (0.2%)	1.05	14/4427 (0.3%)
1	F	1.17	6/3258 (0.2%)	1.07	21/4420 (0.5%)
All	All	1.18	35/19540 (0.2%)	1.04	86/26507 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	236	GLU	CD-OE1	9.28	1.35	1.25
1	D	315	GLU	CD-OE1	8.82	1.35	1.25
1	F	272	GLU	CD-OE1	6.71	1.33	1.25
1	E	151	GLU	CD-OE2	-6.52	1.18	1.25
1	B	309	GLU	CB-CG	6.49	1.64	1.52

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	378	ARG	NE-CZ-NH1	10.78	125.69	120.30
1	B	363	ARG	NE-CZ-NH2	-10.43	115.09	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	242	LEU	CB-CG-CD2	-9.77	94.39	111.00
1	B	111	ARG	NE-CZ-NH1	9.31	124.96	120.30
1	B	363	ARG	NE-CZ-NH1	9.21	124.90	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	165	GLY	Peptide

## 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	3171	0	3105	53	0
1	B	3181	0	3091	65	0
1	C	3180	0	3116	49	0
1	D	3205	0	3134	59	0
1	E	3199	0	3132	52	0
1	F	3186	0	3126	64	0
2	A	43	0	30	6	0
2	B	43	0	30	6	0
2	C	43	0	30	5	0
2	D	43	0	30	6	0
2	E	43	0	30	6	0
2	F	43	0	30	10	0
3	A	77	0	70	16	0
3	B	77	0	70	23	0
3	C	77	0	70	22	0
3	D	77	0	70	10	0
3	E	77	0	67	23	0
3	F	77	0	68	6	0
4	A	10	0	0	8	0
4	B	15	0	0	5	0
4	C	15	0	0	3	0
4	D	15	0	0	5	0
4	E	15	0	0	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	15	0	0	5	0
5	A	4	0	5	0	0
5	B	4	0	5	1	0
5	C	4	0	5	1	0
5	D	4	0	5	0	0
5	E	4	0	5	0	0
5	F	4	0	5	0	0
6	A	448	0	0	17	0
6	B	439	0	0	15	0
6	C	457	0	0	13	0
6	D	431	0	0	12	0
6	E	445	0	0	14	1
6	F	417	0	0	20	1
All	All	22588	0	19329	442	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 11.

The worst 5 of 442 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)
1:E:169:HIS:O	1:E:175:ILE:HD11	1.28	1.29
3:E:1403:BCD:C37	3:E:1403:BCD:O21	1.84	1.25
3:E:1403:BCD:O37	3:E:1403:BCD:H21	1.40	1.21
3:E:1403:BCD:H37	3:E:1403:BCD:O21	1.37	1.19
3:C:1403:BCD:O36	3:C:1403:BCD:H27	1.43	1.15

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)
6:E:2187:HOH:O	6:F:2414:HOH:O[3_545]	1.89	0.31

## 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	405/407 (100%)	393 (97%)	12 (3%)	0	100	100
1	B	407/407 (100%)	394 (97%)	13 (3%)	0	100	100
1	C	407/407 (100%)	391 (96%)	16 (4%)	0	100	100
1	D	407/407 (100%)	395 (97%)	12 (3%)	0	100	100
1	E	406/407 (100%)	395 (97%)	11 (3%)	0	100	100
1	F	406/407 (100%)	397 (98%)	9 (2%)	0	100	100
All	All	2438/2442 (100%)	2365 (97%)	73 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	340/349 (97%)	332 (98%)	8 (2%)	57	46
1	B	341/349 (98%)	330 (97%)	11 (3%)	46	32
1	C	343/349 (98%)	330 (96%)	13 (4%)	40	24
1	D	346/349 (99%)	336 (97%)	10 (3%)	50	36
1	E	345/349 (99%)	333 (96%)	12 (4%)	43	28
1	F	346/349 (99%)	337 (97%)	9 (3%)	54	42
All	All	2061/2094 (98%)	1998 (97%)	63 (3%)	50	33

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	372[A]	ASP
1	D	99[B]	ARG
1	F	105	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	372[B]	ASP
1	C	399	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	323	ASN
1	D	294	HIS
1	F	341	HIS
1	C	341	HIS
1	D	86	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](#)

35 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	HEM	A	1402	1,5	30,50,50	2.50	13 (43%)	24,82,82	2.81	12 (50%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	BCD	A	1403	-	84,84,84	0.77	1 (1%)	126,126,126	2.31	49 (38%)
4	SO4	A	1404	-	4,4,4	0.84	0	6,6,6	1.46	1 (16%)
4	SO4	A	1405	-	4,4,4	0.72	0	6,6,6	0.72	0
5	EDO	A	1406	2	3,3,3	1.32	0	2,2,2	0.93	0
2	HEM	B	1402	1,5	30,50,50	2.54	12 (40%)	24,82,82	2.72	13 (54%)
3	BCD	B	1403	-	84,84,84	0.92	3 (3%)	126,126,126	2.69	59 (46%)
4	SO4	B	1406	-	4,4,4	1.16	0	6,6,6	0.66	0
4	SO4	B	1407	-	4,4,4	0.86	0	6,6,6	0.54	0
4	SO4	B	1408	-	4,4,4	1.16	0	6,6,6	0.95	0
5	EDO	B	1409	2	3,3,3	1.37	0	2,2,2	0.98	0
2	HEM	C	1402	1,5	30,50,50	2.58	9 (30%)	24,82,82	2.55	9 (37%)
3	BCD	C	1403	-	84,84,84	0.99	2 (2%)	126,126,126	2.47	53 (42%)
4	SO4	C	1404	-	4,4,4	1.19	0	6,6,6	0.97	0
4	SO4	C	1405	-	4,4,4	1.08	0	6,6,6	0.35	0
4	SO4	C	1406	-	4,4,4	0.77	0	6,6,6	0.61	0
5	EDO	C	1407	2	3,3,3	0.84	0	2,2,2	0.04	0
2	HEM	D	1402	1,5	30,50,50	2.40	12 (40%)	24,82,82	2.81	12 (50%)
3	BCD	D	1403	-	84,84,84	0.70	1 (1%)	126,126,126	2.03	42 (33%)
4	SO4	D	1404	-	4,4,4	1.12	0	6,6,6	0.59	0
4	SO4	D	1405	-	4,4,4	1.41	0	6,6,6	0.49	0
4	SO4	D	1406	-	4,4,4	0.99	0	6,6,6	1.25	1 (16%)
5	EDO	D	1407	2	3,3,3	1.15	0	2,2,2	0.84	0
2	HEM	E	1402	1,5	30,50,50	2.44	12 (40%)	24,82,82	2.68	11 (45%)
3	BCD	E	1403	-	84,84,84	0.92	0	126,126,126	2.96	55 (43%)
4	SO4	E	1404	-	4,4,4	1.11	0	6,6,6	0.78	0
4	SO4	E	1405	-	4,4,4	1.17	0	6,6,6	0.63	0
4	SO4	E	1406	1	4,4,4	0.85	0	6,6,6	0.83	0
5	EDO	E	1407	2	3,3,3	1.38	0	2,2,2	0.91	0
2	HEM	F	1402	1,5	30,50,50	2.37	7 (23%)	24,82,82	3.05	12 (50%)
3	BCD	F	1403	-	84,84,84	0.69	1 (1%)	126,126,126	2.34	53 (42%)
4	SO4	F	1404	-	4,4,4	0.91	0	6,6,6	0.53	0
4	SO4	F	1405	-	4,4,4	1.57	1 (25%)	6,6,6	1.01	1 (16%)
4	SO4	F	1406	-	4,4,4	0.77	0	6,6,6	0.59	0
5	EDO	F	1407	2	3,3,3	1.30	0	2,2,2	0.73	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
2	HEM	A	1402	1,5	-	0/10/54/54	0/0/8/8
3	BCD	A	1403	-	2/2/35/35	0/42/182/182	0/0/8/8
4	SO4	A	1404	-	-	0/0/0/0	0/0/0/0
4	SO4	A	1405	-	-	0/0/0/0	0/0/0/0
5	EDO	A	1406	2	-	0/1/1/1	0/0/0/0
2	HEM	B	1402	1,5	-	0/10/54/54	0/0/8/8
3	BCD	B	1403	-	2/2/35/35	0/42/182/182	0/0/8/8
4	SO4	B	1406	-	-	0/0/0/0	0/0/0/0
4	SO4	B	1407	-	-	0/0/0/0	0/0/0/0
4	SO4	B	1408	-	-	0/0/0/0	0/0/0/0
5	EDO	B	1409	2	-	0/1/1/1	0/0/0/0
2	HEM	C	1402	1,5	-	0/10/54/54	0/0/8/8
3	BCD	C	1403	-	2/2/35/35	0/42/182/182	0/0/8/8
4	SO4	C	1404	-	-	0/0/0/0	0/0/0/0
4	SO4	C	1405	-	-	0/0/0/0	0/0/0/0
4	SO4	C	1406	-	-	0/0/0/0	0/0/0/0
5	EDO	C	1407	2	-	0/1/1/1	0/0/0/0
2	HEM	D	1402	1,5	-	0/10/54/54	0/0/8/8
3	BCD	D	1403	-	1/1/35/35	0/42/182/182	0/0/8/8
4	SO4	D	1404	-	-	0/0/0/0	0/0/0/0
4	SO4	D	1405	-	-	0/0/0/0	0/0/0/0
4	SO4	D	1406	-	-	0/0/0/0	0/0/0/0
5	EDO	D	1407	2	-	0/1/1/1	0/0/0/0
2	HEM	E	1402	1,5	-	0/10/54/54	0/0/8/8
3	BCD	E	1403	-	3/3/35/35	0/42/182/182	0/0/8/8
4	SO4	E	1404	-	-	0/0/0/0	0/0/0/0
4	SO4	E	1405	-	-	0/0/0/0	0/0/0/0
4	SO4	E	1406	1	-	0/0/0/0	0/0/0/0
5	EDO	E	1407	2	-	0/1/1/1	0/0/0/0
2	HEM	F	1402	1,5	-	0/10/54/54	0/0/8/8
3	BCD	F	1403	-	2/2/35/35	1/42/182/182	0/0/8/8
4	SO4	F	1404	-	-	0/0/0/0	0/0/0/0
4	SO4	F	1405	-	-	0/0/0/0	0/0/0/0
4	SO4	F	1406	-	-	0/0/0/0	0/0/0/0
5	EDO	F	1407	2	-	0/1/1/1	0/0/0/0

The worst 5 of 74 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1402	HEM	C3B-C4B	-9.42	1.43	1.51
2	B	1402	HEM	C3B-C4B	-9.07	1.43	1.51
2	F	1402	HEM	C3B-C4B	-8.91	1.43	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1402	HEM	C3B-C4B	-8.71	1.44	1.51
2	E	1402	HEM	C3B-C4B	-8.49	1.44	1.51

The worst 5 of 383 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	1403	BCD	O54-C54-C44	-10.39	87.79	109.75
3	E	1403	BCD	O55-C55-C45	-9.36	89.97	109.75
3	E	1403	BCD	C15-O44-C44	-8.50	95.79	118.01
3	C	1403	BCD	O53-C53-C43	-8.45	91.90	109.75
3	F	1403	BCD	C37-C47-C57	-8.14	92.44	110.84

5 of 12 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	E	1403	BCD	C11
3	E	1403	BCD	C14
3	E	1403	BCD	C17
3	A	1403	BCD	C11
3	A	1403	BCD	C17

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	F	1403	BCD	C44-O44-C15-C25

There are no ring outliers.

20 monomers are involved in 172 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1402	HEM	6	0
3	A	1403	BCD	16	0
4	A	1404	SO4	8	0
2	B	1402	HEM	6	0
3	B	1403	BCD	23	0
4	B	1408	SO4	5	0
5	B	1409	EDO	1	0
2	C	1402	HEM	5	0
3	C	1403	BCD	22	0
4	C	1406	SO4	3	0
5	C	1407	EDO	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1402	HEM	6	0
3	D	1403	BCD	10	0
4	D	1406	SO4	5	0
2	E	1402	HEM	6	0
3	E	1403	BCD	23	0
4	E	1406	SO4	5	0
2	F	1402	HEM	10	0
3	F	1403	BCD	6	0
4	F	1406	SO4	5	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	402/407 (98%)	-0.30	8 (1%) 68 75	12, 22, 44, 60	0
1	B	404/407 (99%)	-0.29	6 (1%) 76 82	12, 22, 44, 64	0
1	C	402/407 (98%)	-0.27	7 (1%) 73 79	12, 22, 43, 57	0
1	D	402/407 (98%)	-0.25	6 (1%) 76 82	12, 22, 43, 61	0
1	E	402/407 (98%)	-0.27	8 (1%) 68 75	12, 22, 43, 56	0
1	F	402/407 (98%)	-0.27	8 (1%) 68 75	12, 22, 44, 55	0
All	All	2414/2442 (98%)	-0.28	43 (1%) 71 78	12, 22, 44, 64	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	165	GLY	5.5
1	F	170	VAL	4.8
1	C	170	VAL	4.4
1	D	167	SER	4.3
1	B	165	GLY	4.2

### 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 ⓘ

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
4	SO4	C	1406	5/5	0.95	0.57	16.08	32,42,45,47	0
4	SO4	D	1406	5/5	0.96	0.55	10.04	29,44,46,47	0
4	SO4	A	1404	5/5	0.98	0.50	10.00	27,42,43,44	0
5	EDO	B	1409	4/4	0.89	0.15	9.94	19,20,22,32	0
4	SO4	B	1408	5/5	0.95	0.50	8.57	30,43,45,49	0
4	SO4	F	1406	5/5	0.98	0.54	8.55	27,41,41,48	0
4	SO4	E	1406	5/5	0.96	0.50	8.21	29,41,44,49	0
5	EDO	E	1407	4/4	0.95	0.14	5.60	22,23,23,31	0
5	EDO	F	1407	4/4	0.97	0.12	3.55	18,21,22,27	0
5	EDO	D	1407	4/4	0.97	0.11	3.27	19,20,21,34	0
3	BCD	A	1403	77/77	0.89	0.17	2.10	21,40,45,49	77
5	EDO	A	1406	4/4	0.96	0.10	2.04	19,22,23,30	0
3	BCD	F	1403	77/77	0.91	0.15	1.83	26,39,45,48	77
3	BCD	D	1403	77/77	0.91	0.15	1.78	24,39,44,51	77
3	BCD	E	1403	77/77	0.91	0.15	1.68	25,38,42,49	77
3	BCD	B	1403	77/77	0.90	0.15	1.67	21,38,43,46	77
5	EDO	C	1407	4/4	0.98	0.11	1.50	21,23,26,30	0
2	HEM	A	1402	43/43	0.98	0.09	1.14	11,14,17,22	0
2	HEM	D	1402	43/43	0.99	0.09	0.80	8,14,16,22	0
2	HEM	B	1402	43/43	0.99	0.09	0.77	10,13,16,24	0
2	HEM	C	1402	43/43	0.99	0.09	0.77	9,14,17,24	0
2	HEM	E	1402	43/43	0.99	0.09	0.74	9,14,17,22	0
2	HEM	F	1402	43/43	0.98	0.09	0.72	10,14,19,28	0
3	BCD	C	1403	77/77	0.90	0.13	0.66	22,37,45,50	77
4	SO4	F	1404	5/5	0.98	0.16	-	25,25,31,32	0
4	SO4	C	1404	5/5	0.98	0.12	-	26,28,29,32	0
4	SO4	B	1406	5/5	0.97	0.14	-	30,30,35,36	0
4	SO4	B	1407	5/5	0.98	0.12	-	23,29,30,31	0
4	SO4	A	1405	5/5	0.98	0.14	-	27,30,31,33	0
4	SO4	E	1405	5/5	0.97	0.20	-	34,34,35,38	0
4	SO4	D	1405	5/5	0.95	0.18	-	30,31,35,35	0
4	SO4	D	1404	5/5	0.98	0.14	-	25,26,31,31	0
4	SO4	E	1404	5/5	0.97	0.15	-	27,28,31,32	0
4	SO4	F	1405	5/5	0.96	0.21	-	33,34,36,37	0
4	SO4	C	1405	5/5	0.97	0.17	-	33,33,35,36	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.