



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 05:31 am GMT

PDB ID : 3ABA  
Title : Crystal structure of CYP105P1 in complex with filipin I  
Authors : Xu, L.H.; Fushinobu, S.; Takamatsu, S.; Wakagi, T.; Ikeda, H.; Shoun, H.  
Deposited on : 2009-12-04  
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 : trunk28620  
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) : recalc28949

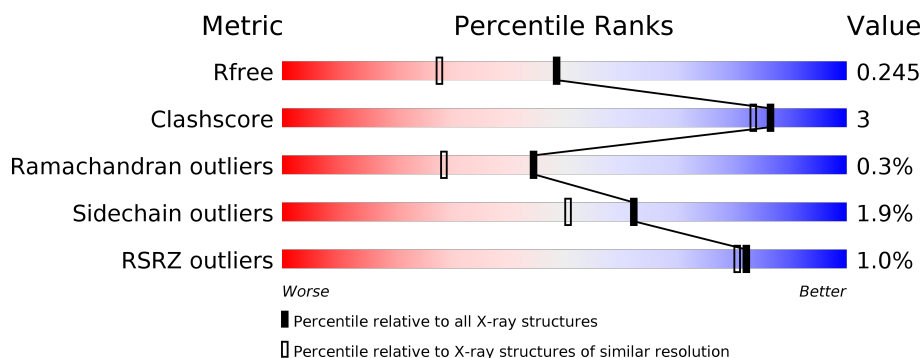
# 1 Overall quality at a glance

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	403	<div> <div></div> <div>88%</div> <div>9%</div> <div>..</div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3775 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 P450.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	397	3081	1943	555	571	12	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

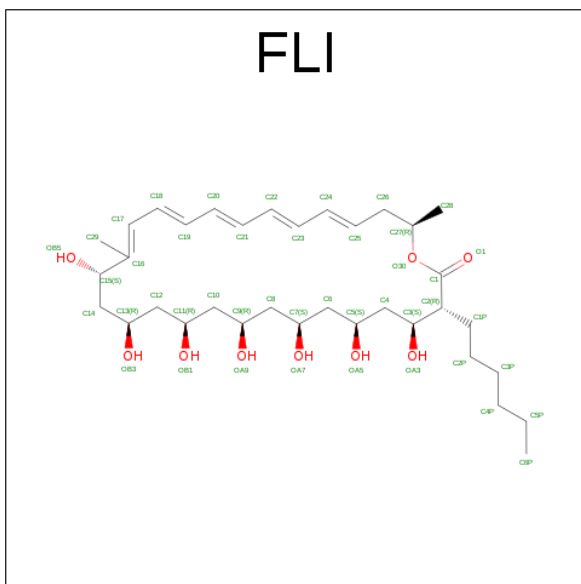
Chain	Residue	Modelled	Actual	Comment	Reference
A	400	HIS	-	EXPRESSION TAG	UNP Q93H81
A	401	HIS	-	EXPRESSION TAG	UNP Q93H81
A	402	HIS	-	EXPRESSION TAG	UNP Q93H81
A	403	HIS	-	EXPRESSION TAG	UNP Q93H81

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



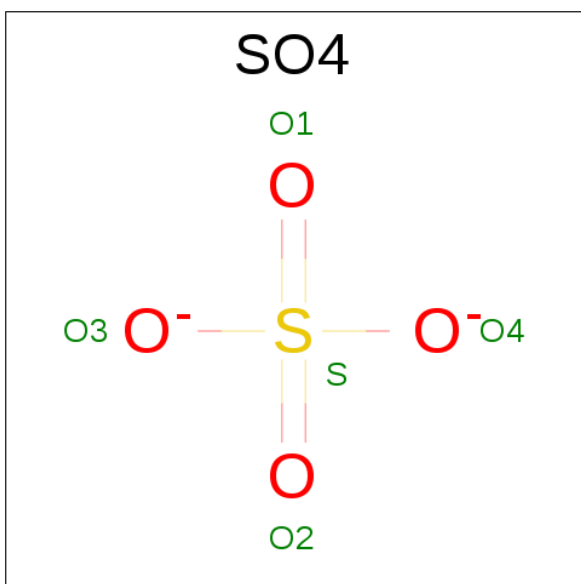
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
2	A	1	43	34	1	4	4	0	0

- Molecule 3 is (3R,4S,6S,8S,10R,12R,14R,16S,17E,19E,21E,23E,25E,28R)-3-HEXYL-4,6,8,10,12,14,16-HEPTAHYDROXY-17,28-DIMETHYLOXACYCLOOCTACOSA-17,19,21,23,25-PENTAEN-2-ONE (three-letter code: FLI) (formula: C<sub>35</sub>H<sub>58</sub>O<sub>9</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			44	35	9		

- 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		

*Continued on next page...*

*Continued from previous page...*

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		

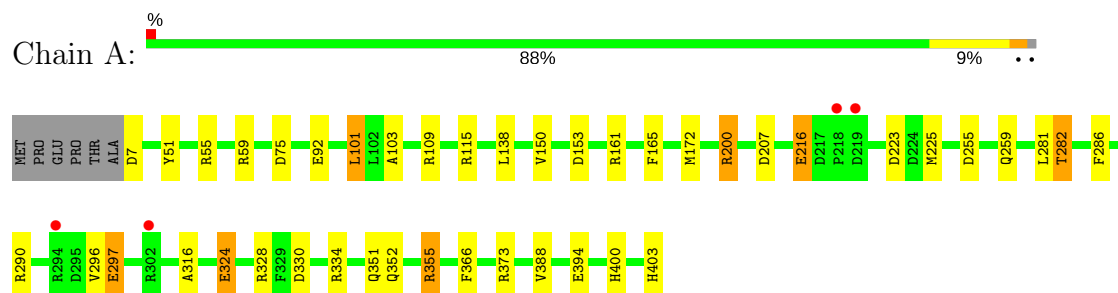
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	592	Total	O	0	0
			592	592		

### 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 P450



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.37Å 91.37Å 151.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.74 – 1.80 39.74 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.2 (39.74-1.80) 98.2 (39.74-1.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.73 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.188 , 0.238 0.198 , 0.245	Depositor DCC
$R_{free}$ test set	2980 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.1	Xtriage
Anisotropy	0.017	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 33.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3775	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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: FLI, HEM, SO4

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	1.34	13/3147 (0.4%)	1.12	17/4279 (0.4%)

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	394	GLU	CD-OE1	8.44	1.34	1.25
1	A	403	HIS	C-O	7.03	1.36	1.23
1	A	388	VAL	CB-CG2	6.72	1.67	1.52
1	A	286	PHE	CE1-CZ	6.71	1.50	1.37
1	A	115	ARG	CZ-NH1	6.63	1.41	1.33
1	A	92	GLU	CG-CD	6.39	1.61	1.51
1	A	200	ARG	CZ-NH1	5.58	1.40	1.33
1	A	103	ALA	CA-CB	5.42	1.63	1.52
1	A	334	ARG	CG-CD	5.30	1.65	1.51
1	A	150	VAL	CB-CG2	5.23	1.63	1.52
1	A	394	GLU	CB-CG	-5.05	1.42	1.52
1	A	165	PHE	CE2-CZ	5.05	1.47	1.37
1	A	366	PHE	CE1-CZ	5.02	1.46	1.37

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	200	ARG	NE-CZ-NH2	-13.77	113.42	120.30
1	A	109	ARG	NE-CZ-NH1	12.25	126.42	120.30
1	A	109	ARG	NE-CZ-NH2	-9.23	115.69	120.30
1	A	355	ARG	NE-CZ-NH2	-8.13	116.23	120.30
1	A	403	HIS	CA-C-O	-7.74	103.84	120.10
1	A	115	ARG	NE-CZ-NH2	-7.20	116.70	120.30
1	A	172	MET	CG-SD-CE	-6.53	89.76	100.20
1	A	355	ARG	NE-CZ-NH1	6.47	123.54	120.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	161	ARG	NE-CZ-NH2	-6.37	117.11	120.30
1	A	153	ASP	CB-CG-OD1	6.25	123.92	118.30
1	A	101	LEU	CA-CB-CG	6.10	129.33	115.30
1	A	330	ASP	CB-CG-OD1	5.75	123.47	118.30
1	A	138	LEU	CB-CG-CD1	-5.67	101.37	111.00
1	A	223	ASP	CB-CG-OD1	5.39	123.15	118.30
1	A	109	ARG	CD-NE-CZ	5.25	130.95	123.60
1	A	75	ASP	CB-CG-OD1	5.24	123.02	118.30
1	A	290	ARG	NE-CZ-NH1	5.01	122.81	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	3081	0	3055	16	0
2	A	43	0	30	0	0
3	A	44	0	58	0	0
4	A	15	0	0	1	0
5	A	592	0	0	6	0
All	All	3775	0	3143	16	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 3.

All (16) 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:A:59:ARG:HD3	5:A:965:HOH:O	1.91	0.70
1:A:51:TYR:CE2	1:A:55:ARG:HD2	2.44	0.53
1:A:51:TYR:HA	1:A:316:ALA:HB1	1.92	0.52
1:A:259:GLN:HG2	4:A:406:SO4:O1	2.13	0.48
1:A:324:GLU:OE2	1:A:328:ARG:NH2	2.40	0.47
1:A:225:MET:HE2	5:A:445:HOH:O	2.15	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:351:GLN:OE1	1:A:355:ARG:NH1	2.46	0.43
1:A:225:MET:HG3	5:A:521:HOH:O	2.19	0.43
1:A:297:GLU:HG2	5:A:423:HOH:O	2.18	0.43
1:A:281:LEU:O	1:A:282:THR:C	2.58	0.42
1:A:373:ARG:HH22	1:A:400:HIS:CD2	2.37	0.42
1:A:200:ARG:HD3	1:A:207:ASP:OD1	2.20	0.41
1:A:216:GLU:HB2	5:A:878:HOH:O	2.18	0.41
1:A:255:ASP:OD1	5:A:468:HOH:O	2.22	0.41
1:A:200:ARG:HH21	1:A:200:ARG:HD2	1.62	0.40
1:A:352:GLN:NE2	1:A:355:ARG:HH11	2.19	0.40

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	395/403 (98%)	388 (98%)	6 (2%)	1 (0%)	44 29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	282	THR

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	323/328 (98%)	317 (98%)	6 (2%)	62 50

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	ASP
1	A	101	LEU
1	A	216	GLU
1	A	296	VAL
1	A	297	GLU
1	A	324	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	GLN
1	A	272	ASN
1	A	352	GLN
1	A	400	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

5 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	1408	1	28,50,50	2.24	8 (28%)	17,82,82	2.84	7 (41%)
3	FLI	A	404	-	44,44,44	2.10	13 (29%)	49,55,55	2.19	17 (34%)
4	SO4	A	405	-	4,4,4	0.18	0	6,6,6	0.92	0
4	SO4	A	406	-	4,4,4	0.47	0	6,6,6	1.50	1 (16%)
4	SO4	A	407	-	4,4,4	0.66	0	6,6,6	0.59	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	1408	1	-	0/6/54/54	0/0/8/8
3	FLI	A	404	-	-	0/59/59/59	0/0/1/1
4	SO4	A	405	-	-	0/0/0/0	0/0/0/0
4	SO4	A	406	-	-	0/0/0/0	0/0/0/0
4	SO4	A	407	-	-	0/0/0/0	0/0/0/0

All (21) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1408	HEM	C3C-C2C	-6.30	1.32	1.40
2	A	1408	HEM	C3B-C2B	-4.84	1.34	1.40
2	A	1408	HEM	C3C-CAC	2.04	1.51	1.47
3	A	404	FLI	C4-C5	2.06	1.56	1.52
3	A	404	FLI	OB5-C15	2.23	1.48	1.42
3	A	404	FLI	C15-C16	2.26	1.54	1.51
3	A	404	FLI	C4-C3	2.28	1.58	1.53
2	A	1408	HEM	CMB-C2B	2.36	1.56	1.51
3	A	404	FLI	C10-C9	2.41	1.57	1.52
3	A	404	FLI	C14-C13	2.44	1.57	1.52
3	A	404	FLI	C24-C25	2.53	1.41	1.32
3	A	404	FLI	C26-C27	2.85	1.55	1.52
3	A	404	FLI	C21-C20	2.87	1.43	1.35
3	A	404	FLI	C14-C15	2.96	1.58	1.53
2	A	1408	HEM	C3B-CAB	3.17	1.54	1.47
2	A	1408	HEM	C1C-NC	3.28	1.40	1.36
2	A	1408	HEM	CMC-C2C	3.36	1.58	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	404	FLI	C10-C11	4.03	1.60	1.52
3	A	404	FLI	C29-C16	4.26	1.58	1.50
2	A	1408	HEM	C3D-C2D	4.34	1.50	1.37
3	A	404	FLI	C17-C16	7.23	1.40	1.33

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	404	FLI	C27-C26-C25	-4.73	104.98	113.88
2	A	1408	HEM	CMA-C3A-C4A	-3.88	122.50	128.46
3	A	404	FLI	C11-C10-C9	-3.85	108.92	114.22
3	A	404	FLI	C24-C23-C22	-3.47	117.08	124.88
3	A	404	FLI	C7-C6-C5	-3.40	109.53	114.22
3	A	404	FLI	C9-C8-C7	-3.37	109.57	114.22
3	A	404	FLI	C14-C15-C16	-3.35	106.18	112.28
2	A	1408	HEM	CBD-CAD-C3D	-3.06	106.63	112.47
3	A	404	FLI	C8-C7-C6	-3.05	107.49	112.56
3	A	404	FLI	OB5-C15-C14	-3.01	100.98	108.61
3	A	404	FLI	C22-C21-C20	-2.98	118.19	124.88
3	A	404	FLI	O30-C27-C26	-2.93	102.09	107.21
3	A	404	FLI	C26-C25-C24	-2.71	121.65	125.40
3	A	404	FLI	O30-C1-O1	-2.65	118.71	123.90
2	A	1408	HEM	CMD-C2D-C1D	-2.21	125.07	128.46
3	A	404	FLI	C10-C9-C8	-2.21	108.88	112.56
4	A	406	SO4	O4-S-O3	-2.16	99.23	108.96
3	A	404	FLI	C6-C5-C4	-2.12	109.03	112.56
3	A	404	FLI	OB3-C13-C14	2.19	114.09	109.14
2	A	1408	HEM	CMD-C2D-C3D	2.32	129.31	124.94
2	A	1408	HEM	CMC-C2C-C3C	2.58	129.68	124.89
2	A	1408	HEM	CMA-C3A-C2A	2.69	130.02	124.94
3	A	404	FLI	C13-C14-C15	3.46	119.49	114.31
3	A	404	FLI	O30-C1-C2	6.63	120.05	111.26
2	A	1408	HEM	C4C-C3C-C2C	8.39	112.76	106.90

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	406	SO4	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 [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	397/403 (98%)	-0.23	4 (1%) 82 80	10, 17, 29, 41	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	294	ARG	2.7
1	A	302	ARG	2.6
1	A	219	ASP	2.2
1	A	218	PRO	2.2

### 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 [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
4	SO4	A	405	5/5	0.98	0.10	0.89	41,42,44,47	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
2	HEM	A	1408	43/43	0.98	0.13	0.49	17,20,26,32	0
3	FLI	A	404	44/44	0.94	0.12	0.28	21,25,30,31	0
4	SO4	A	406	5/5	0.97	0.11	-	39,42,42,44	0
4	SO4	A	407	5/5	0.94	0.26	-	45,46,55,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.