



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:05 PM GMT

PDB ID : 1IO8  
Title : Thermophilic cytochrome P450 (CYP119) from sulfolobus solfataricus: High resolution structural origin of its thermostability and functional properties  
Authors : Park, S.-Y.; Yamane, K.; Adachi, S.; Shiro, Y.; Sligar, S.G.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)  
Deposited on : 2001-02-08  
Resolution : 2.00 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

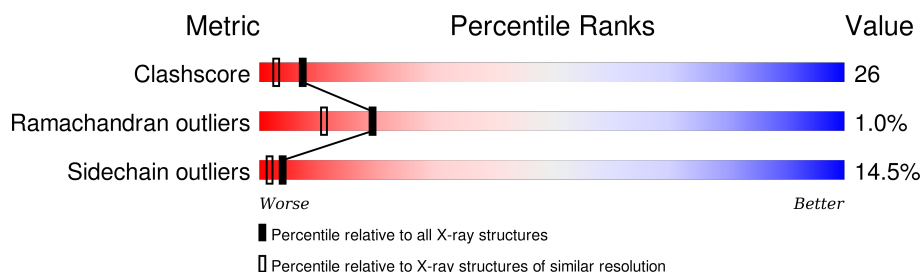
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	368	
1	B	368	



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 CYP119.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	366	Total 3007	C 1925	N 517	O 559	S 6	0	0	0
1	B	359	Total 2958	C 1893	N 509	O 550	S 6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	LEU	PHE	ENGINEERED	UNP Q55080
B	524	LEU	PHE	ENGINEERED	UNP Q55080

- 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
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

*Continued on next page...*

*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		

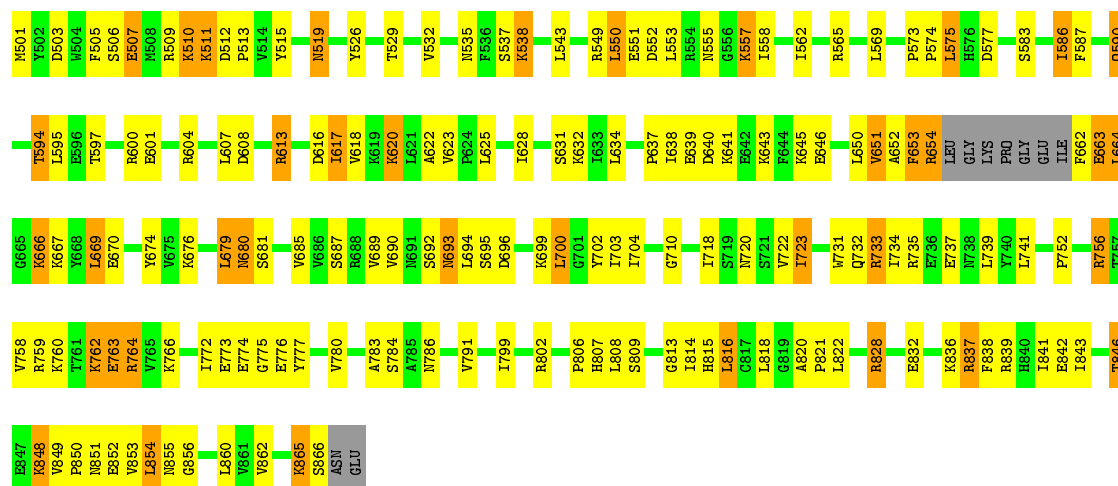
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	132	Total	O	0	0
			132	132		
3	B	151	Total	O	0	0
			151	151		



Note EDS was not executed.

- Chain A:  54% 38% 7%



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.52Å 85.52Å 221.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00	Depositor
% Data completeness (in resolution range)	94.5 (20.00-2.00)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.230 , 0.307	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6334	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: HEM

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	0.41	0/3070	0.65	0/4149
1	B	0.44	0/3019	0.64	0/4079
All	All	0.43	0/6089	0.64	0/8228

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3007	0	3043	156	0
1	B	2958	0	2985	157	0
2	A	43	0	30	3	0
2	B	43	0	30	1	0
3	A	132	0	0	7	0
3	B	151	0	0	15	0
All	All	6334	0	6088	312	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 26.

The worst 5 of 312 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:B:851:ASN:ND2	1:B:853:VAL:H	1.60	0.99
1:B:613:ARG:HB3	1:B:613:ARG:HH11	1.26	0.98
1:B:764:ARG:HB2	1:B:764:ARG:HH11	1.24	0.98
1:B:586:ILE:HD11	1:B:818:LEU:HD21	1.47	0.97
1:A:223:ILE:HD11	1:A:360:LEU:HD11	1.48	0.96

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	364/368 (99%)	338 (93%)	22 (6%)	4 (1%)	17	9
1	B	355/368 (96%)	333 (94%)	19 (5%)	3 (1%)	24	15
All	All	719/736 (98%)	671 (93%)	41 (6%)	7 (1%)	19	11

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	114	GLU
1	B	653	PHE
1	A	139	GLU
1	B	513	PRO
1	A	137	PRO

### 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	336/338 (99%)	290 (86%)	46 (14%)	4	2
1	B	331/338 (98%)	280 (85%)	51 (15%)	3	1
All	All	667/676 (99%)	570 (86%)	97 (14%)	4	2

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	349	VAL
1	B	557	LYS
1	B	816	LEU
1	A	366	SER
1	B	511	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	691	ASN
1	B	693	ASN
1	B	807	HIS
1	B	680	ASN
1	B	786	ASN

### 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

2 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	401	1,3	30,50,50	2.78	8 (26%)	24,82,82	1.94	6 (25%)
2	HEM	B	901	1,3	30,50,50	2.71	9 (30%)	24,82,82	2.25	9 (37%)

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	401	1,3	-	0/10/54/54	0/0/8/8
2	HEM	B	901	1,3	-	0/10/54/54	0/0/8/8

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	HEM	C3C-CAC	-6.25	1.39	1.51
2	A	401	HEM	C3B-CAB	-6.18	1.39	1.51
2	B	901	HEM	C3B-CAB	-5.88	1.40	1.51
2	B	901	HEM	C3C-CAC	-5.42	1.41	1.51
2	A	401	HEM	C2D-C3D	-5.42	1.38	1.54

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	901	HEM	CAA-C2A-C1A	-3.37	123.35	127.01
2	B	901	HEM	CBD-CAD-C3D	-2.96	104.93	113.55
2	B	901	HEM	C3C-CAC-CBC	-2.88	120.03	124.46
2	A	401	HEM	CMC-C2C-C3C	2.38	122.47	116.53
2	B	901	HEM	CMC-C2C-C3C	2.67	123.21	116.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	HEM	3	0
2	B	901	HEM	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.