



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:47 PM GMT

PDB ID : 4PXH  
Title : Structure of P450sky (CYP163B3), a cytochrome P450 from skyllamycin biosynthesis in complex with a peptidyl carrier protein domain  
Authors : Haslinger, K.; Cryle, M.J.  
Deposited on : 2014-03-24  
Resolution : 2.70 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 (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

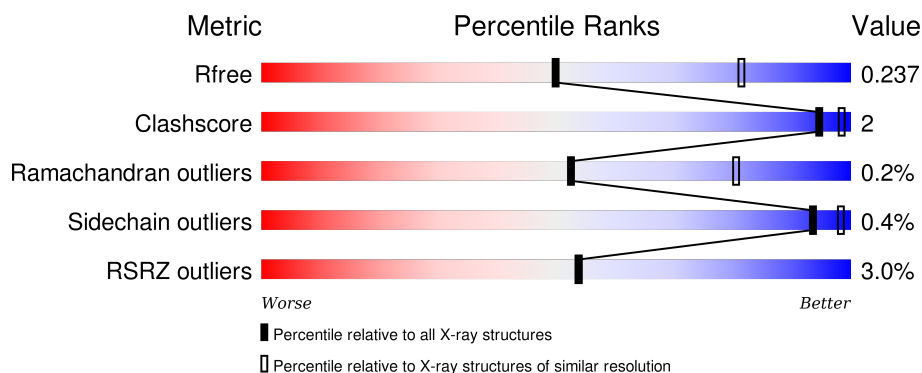
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	446	<div> <div>2%</div> <div>90%</div> <div>7%</div> </div>
1	C	446	<div> <div>%</div> <div>86%</div> <div>5%</div> <div>8%</div> </div>
1	E	446	<div> <div>3%</div> <div>87%</div> <div>9%</div> </div>
2	B	90	<div> <div>11%</div> <div>81%</div> <div>17%</div> </div>
2	D	90	<div> <div>4%</div> <div>80%</div> <div>18%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	90	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11743 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 monooxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	417	Total	C	N	O	S	0	1	0
			3262	2055	568	629	10			
1	C	409	Total	C	N	O	S	0	1	0
			3206	2020	559	617	10			
1	E	407	Total	C	N	O	S	0	0	0
			3182	2007	553	612	10			

There are 69 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	EXPRESSION TAG	UNP F2YRY7
A	-21	GLY	-	EXPRESSION TAG	UNP F2YRY7
A	-20	SER	-	EXPRESSION TAG	UNP F2YRY7
A	-19	SER	-	EXPRESSION TAG	UNP F2YRY7
A	-18	HIS	-	EXPRESSION TAG	UNP F2YRY7
A	-17	HIS	-	EXPRESSION TAG	UNP F2YRY7
A	-16	HIS	-	EXPRESSION TAG	UNP F2YRY7
A	-15	HIS	-	EXPRESSION TAG	UNP F2YRY7
A	-14	HIS	-	EXPRESSION TAG	UNP F2YRY7
A	-13	HIS	-	EXPRESSION TAG	UNP F2YRY7
A	-12	SER	-	EXPRESSION TAG	UNP F2YRY7
A	-11	SER	-	EXPRESSION TAG	UNP F2YRY7
A	-10	GLY	-	EXPRESSION TAG	UNP F2YRY7
A	-9	LEU	-	EXPRESSION TAG	UNP F2YRY7
A	-8	VAL	-	EXPRESSION TAG	UNP F2YRY7
A	-7	PRO	-	EXPRESSION TAG	UNP F2YRY7
A	-6	ARG	-	EXPRESSION TAG	UNP F2YRY7
A	-5	GLY	-	EXPRESSION TAG	UNP F2YRY7
A	-4	SER	-	EXPRESSION TAG	UNP F2YRY7
A	-3	HIS	-	EXPRESSION TAG	UNP F2YRY7
A	-2	SER	-	EXPRESSION TAG	UNP F2YRY7
A	-1	GLY	-	EXPRESSION TAG	UNP F2YRY7
A	0	HIS	-	EXPRESSION TAG	UNP F2YRY7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-22	MET	-	EXPRESSION TAG	UNP F2YRY7
C	-21	GLY	-	EXPRESSION TAG	UNP F2YRY7
C	-20	SER	-	EXPRESSION TAG	UNP F2YRY7
C	-19	SER	-	EXPRESSION TAG	UNP F2YRY7
C	-18	HIS	-	EXPRESSION TAG	UNP F2YRY7
C	-17	HIS	-	EXPRESSION TAG	UNP F2YRY7
C	-16	HIS	-	EXPRESSION TAG	UNP F2YRY7
C	-15	HIS	-	EXPRESSION TAG	UNP F2YRY7
C	-14	HIS	-	EXPRESSION TAG	UNP F2YRY7
C	-13	HIS	-	EXPRESSION TAG	UNP F2YRY7
C	-12	SER	-	EXPRESSION TAG	UNP F2YRY7
C	-11	SER	-	EXPRESSION TAG	UNP F2YRY7
C	-10	GLY	-	EXPRESSION TAG	UNP F2YRY7
C	-9	LEU	-	EXPRESSION TAG	UNP F2YRY7
C	-8	VAL	-	EXPRESSION TAG	UNP F2YRY7
C	-7	PRO	-	EXPRESSION TAG	UNP F2YRY7
C	-6	ARG	-	EXPRESSION TAG	UNP F2YRY7
C	-5	GLY	-	EXPRESSION TAG	UNP F2YRY7
C	-4	SER	-	EXPRESSION TAG	UNP F2YRY7
C	-3	HIS	-	EXPRESSION TAG	UNP F2YRY7
C	-2	SER	-	EXPRESSION TAG	UNP F2YRY7
C	-1	GLY	-	EXPRESSION TAG	UNP F2YRY7
C	0	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-22	MET	-	EXPRESSION TAG	UNP F2YRY7
E	-21	GLY	-	EXPRESSION TAG	UNP F2YRY7
E	-20	SER	-	EXPRESSION TAG	UNP F2YRY7
E	-19	SER	-	EXPRESSION TAG	UNP F2YRY7
E	-18	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-17	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-16	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-15	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-14	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-13	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-12	SER	-	EXPRESSION TAG	UNP F2YRY7
E	-11	SER	-	EXPRESSION TAG	UNP F2YRY7
E	-10	GLY	-	EXPRESSION TAG	UNP F2YRY7
E	-9	LEU	-	EXPRESSION TAG	UNP F2YRY7
E	-8	VAL	-	EXPRESSION TAG	UNP F2YRY7
E	-7	PRO	-	EXPRESSION TAG	UNP F2YRY7
E	-6	ARG	-	EXPRESSION TAG	UNP F2YRY7
E	-5	GLY	-	EXPRESSION TAG	UNP F2YRY7
E	-4	SER	-	EXPRESSION TAG	UNP F2YRY7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-3	HIS	-	EXPRESSION TAG	UNP F2YRY7
E	-2	SER	-	EXPRESSION TAG	UNP F2YRY7
E	-1	GLY	-	EXPRESSION TAG	UNP F2YRY7
E	0	HIS	-	EXPRESSION TAG	UNP F2YRY7

- Molecule 2 is a protein called Peptide synthetase.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	75	Total	C	N	O	0	0	0
			605	379	110	116			
2	D	74	Total	C	N	O	0	0	0
			601	377	109	115			
2	F	74	Total	C	N	O	0	0	0
			601	377	109	115			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	82	ALA	-	EXPRESSION TAG	UNP F2YRY5
B	83	ALA	-	EXPRESSION TAG	UNP F2YRY5
B	84	LEU	-	EXPRESSION TAG	UNP F2YRY5
B	85	GLU	-	EXPRESSION TAG	UNP F2YRY5
B	86	HIS	-	EXPRESSION TAG	UNP F2YRY5
B	87	HIS	-	EXPRESSION TAG	UNP F2YRY5
B	88	HIS	-	EXPRESSION TAG	UNP F2YRY5
B	89	HIS	-	EXPRESSION TAG	UNP F2YRY5
B	90	HIS	-	EXPRESSION TAG	UNP F2YRY5
B	91	HIS	-	EXPRESSION TAG	UNP F2YRY5
D	82	ALA	-	EXPRESSION TAG	UNP F2YRY5
D	83	ALA	-	EXPRESSION TAG	UNP F2YRY5
D	84	LEU	-	EXPRESSION TAG	UNP F2YRY5
D	85	GLU	-	EXPRESSION TAG	UNP F2YRY5
D	86	HIS	-	EXPRESSION TAG	UNP F2YRY5
D	87	HIS	-	EXPRESSION TAG	UNP F2YRY5
D	88	HIS	-	EXPRESSION TAG	UNP F2YRY5
D	89	HIS	-	EXPRESSION TAG	UNP F2YRY5
D	90	HIS	-	EXPRESSION TAG	UNP F2YRY5
D	91	HIS	-	EXPRESSION TAG	UNP F2YRY5
F	82	ALA	-	EXPRESSION TAG	UNP F2YRY5
F	83	ALA	-	EXPRESSION TAG	UNP F2YRY5
F	84	LEU	-	EXPRESSION TAG	UNP F2YRY5
F	85	GLU	-	EXPRESSION TAG	UNP F2YRY5

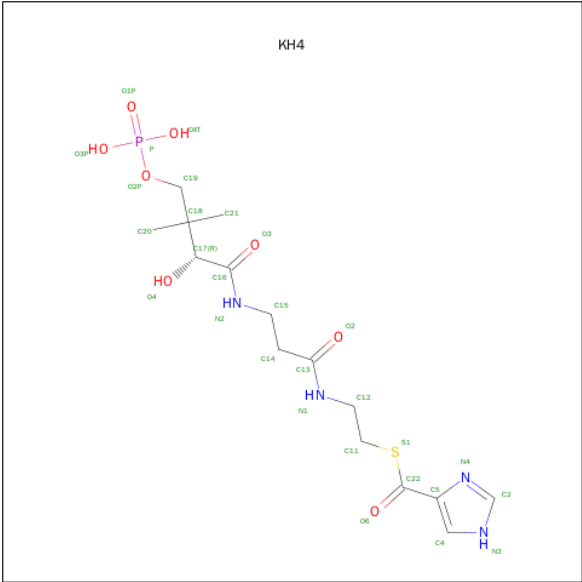
*Continued on next page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	86	HIS	-	EXPRESSION TAG	UNP F2YRY5
F	87	HIS	-	EXPRESSION TAG	UNP F2YRY5
F	88	HIS	-	EXPRESSION TAG	UNP F2YRY5
F	89	HIS	-	EXPRESSION TAG	UNP F2YRY5
F	90	HIS	-	EXPRESSION TAG	UNP F2YRY5
F	91	HIS	-	EXPRESSION TAG	UNP F2YRY5

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

- Molecule 4 is S-[2-({N-[(2R)-2-HYDROXY-3,3-DIMETHYL-4-(PHOSPHONOOXY)BUTANOYL]-BETA-ALANYL}AMINO)ETHYL] 1H-IMIDAZOLE-4-CARBOETHIOATE (three-letter code: KH4) (formula: C<sub>15</sub>H<sub>25</sub>N<sub>4</sub>O<sub>8</sub>PS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
4	B	1	Total 28	C 15	N 4	O 7	P 1	S 1	0	0
4	D	1	Total 28	C 15	N 4	O 7	P 1	S 1	0	0
4	F	1	Total 28	C 15	N 4	O 7	P 1	S 1	0	0

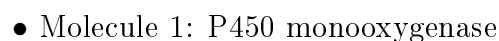
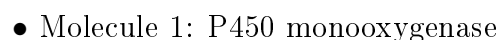
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	19	Total	O	0	0
			19	19		
5	B	2	Total	O	0	0
			2	2		
5	C	34	Total	O	0	0
			34	34		
5	D	3	Total	O	0	0
			3	3		
5	E	15	Total	O	0	0
			15	15		

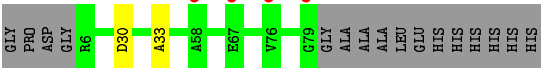
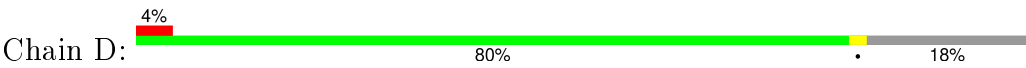




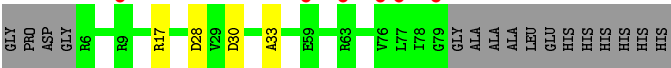
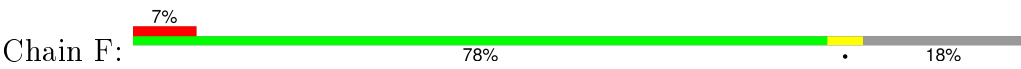
- Molecule 1: P450 monooxygenase



● Molecule 2: Peptide synthetase



● Molecule 2: Peptide synthetase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	161.70 Å   161.70 Å   337.49 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	47.89 – 2.70 47.89 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.89-2.70) 100.0 (47.89-2.70)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.47 (at 2.69 Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.214   ,   0.246 0.210   ,   0.237	Depositor DCC
$R_{free}$ test set	3577 reflections (5.22%)	DCC
Wilson B-factor (Å <sup>2</sup> )	51.4	Xtriage
Anisotropy	0.066	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 30.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 72096 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11743	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.08% 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: HEM, KH4

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	0.32	0/3333	0.47	1/4540 (0.0%)
1	C	0.32	0/3275	0.47	0/4460
1	E	0.32	0/3248	0.46	0/4424
2	B	0.34	0/616	0.42	0/830
2	D	0.34	0/612	0.43	0/825
2	F	0.33	0/612	0.44	0/825
All	All	0.32	0/11696	0.46	1/15904 (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	358	LEU	CA-CB-CG	5.03	126.87	115.30

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	3262	0	3214	10	0
1	C	3206	0	3159	17	0
1	E	3182	0	3136	11	0
2	B	605	0	584	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	601	0	581	1	0
2	F	601	0	581	1	0
3	A	43	0	30	0	0
3	C	43	0	30	0	0
3	E	43	0	30	0	0
4	B	28	0	23	4	0
4	D	28	0	23	4	0
4	F	28	0	23	2	0
5	A	19	0	0	0	0
5	B	2	0	0	0	0
5	C	34	0	0	1	0
5	D	3	0	0	0	0
5	E	15	0	0	0	0
All	All	11743	0	11414	40	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 2.

All (40) 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:C:180:LEU:HD21	4:D:101:KH4:H10	1.66	0.77
1:A:180:LEU:HD21	4:B:101:KH4:H10	1.80	0.62
1:C:76:LYS:HB2	1:C:300:ARG:HG3	1.81	0.62
1:E:180:LEU:HD21	4:F:101:KH4:H10	1.81	0.61
2:F:30:ASP:HB3	2:F:33:ALA:HB2	1.85	0.59
1:E:256:MET:CE	1:E:366:GLU:HB3	2.38	0.54
1:C:374:LEU:O	1:C:378:THR:HB	2.07	0.53
1:E:110:LEU:HD11	1:E:222:LEU:HD21	1.92	0.52
1:A:136:ASP:OD2	1:C:132:ARG:NH2	2.36	0.52
1:C:329:GLU:CD	1:C:329:GLU:H	2.13	0.51
1:A:153:ILE:HB	1:A:154:PRO:HD3	1.93	0.51
1:C:285:GLU:OE2	1:C:364:ARG:HD3	2.11	0.50
1:E:24:GLU:HG2	1:E:389:ARG:HE	1.76	0.49
1:A:374:LEU:O	1:A:378:THR:HB	2.13	0.48
1:E:59:TYR:CD1	1:E:331:VAL:HG21	2.48	0.48
1:A:180:LEU:HD11	4:B:101:KH4:H5	1.95	0.48
1:C:153:ILE:HB	1:C:154:PRO:HD3	1.97	0.47
2:D:30:ASP:HB3	2:D:33:ALA:HB2	1.97	0.46
1:E:153:ILE:HB	1:E:154:PRO:HD3	1.98	0.46
1:C:155:LEU:HD22	1:C:173:LEU:HB2	1.97	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:234:GLU:O	1:A:238:VAL:HG23	2.15	0.45
1:C:123:VAL:HA	1:C:126:ILE:HD12	1.98	0.45
1:C:247:GLY:HA2	4:D:101:KH4:N4	2.32	0.44
1:E:110:LEU:CD1	1:E:222:LEU:HD21	2.47	0.44
1:A:180:LEU:HD11	4:B:101:KH4:N1	2.33	0.44
1:E:385:GLY:H	1:E:402:PRO:HB2	1.82	0.43
1:C:180:LEU:HD11	4:D:101:KH4:H5	1.99	0.43
2:B:48:LEU:HD23	2:B:65:PHE:CE1	2.53	0.43
1:C:274:ARG:HD2	1:C:275:ASP:OD1	2.18	0.43
1:E:65:LEU:HD22	1:E:72:LEU:HD11	2.01	0.42
1:C:187:GLN:NE2	5:C:612:HOH:O	2.52	0.42
1:C:76:LYS:CB	1:C:300:ARG:HG3	2.48	0.42
1:E:254:LEU:HD22	1:E:397:GLY:HA3	2.00	0.42
1:A:180:LEU:CD1	4:B:101:KH4:H5	2.49	0.42
1:A:285:GLU:OE2	1:A:364:ARG:HD3	2.19	0.42
1:C:41:LEU:HD23	1:C:310:GLN:HB2	2.02	0.41
1:C:180:LEU:HD11	4:D:101:KH4:N1	2.34	0.41
1:A:59:TYR:CD1	1:A:331:VAL:HG21	2.56	0.40
1:E:180:LEU:CD1	4:F:101:KH4:H5	2.52	0.40
1:C:59:TYR:CD1	1:C:331:VAL:HG21	2.57	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	416/446 (93%)	405 (97%)	10 (2%)	1 (0%)	52 80
1	C	408/446 (92%)	399 (98%)	8 (2%)	1 (0%)	52 80
1	E	405/446 (91%)	394 (97%)	10 (2%)	1 (0%)	52 80
2	B	73/90 (81%)	72 (99%)	1 (1%)	0	100 100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	72/90 (80%)	71 (99%)	1 (1%)	0	100	100
2	F	72/90 (80%)	72 (100%)	0	0	100	100
All	All	1446/1608 (90%)	1413 (98%)	30 (2%)	3 (0%)	52	80

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	153	ILE
1	A	153	ILE
1	C	153	ILE

### 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	352/375 (94%)	351 (100%)	1 (0%)	94	99
1	C	345/375 (92%)	344 (100%)	1 (0%)	94	99
1	E	342/375 (91%)	341 (100%)	1 (0%)	94	99
2	B	65/75 (87%)	65 (100%)	0	100	100
2	D	65/75 (87%)	65 (100%)	0	100	100
2	F	65/75 (87%)	63 (97%)	2 (3%)	47	78
All	All	1234/1350 (91%)	1229 (100%)	5 (0%)	93	98

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	418	ASP
1	C	305	MET
1	E	329	GLU
2	F	17	ARG
2	F	28	ASP

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	197	ASN
1	C	34	GLN
2	F	41	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 ⓘ

6 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$
3	HEM	A	501	1,4	30,50,50	2.17	10 (33%)	24,82,82	2.38	12 (50%)
4	KH4	B	101	3,2	19,28,29	2.01	2 (10%)	25,37,40	1.51	6 (24%)
3	HEM	C	501	1,4	30,50,50	2.20	9 (30%)	24,82,82	2.36	12 (50%)
4	KH4	D	101	3,2	19,28,29	1.94	2 (10%)	25,37,40	1.39	3 (12%)
3	HEM	E	501	1,4	30,50,50	2.17	9 (30%)	24,82,82	2.40	12 (50%)
4	KH4	F	101	3,2	19,28,29	1.98	2 (10%)	25,37,40	2.80	4 (16%)

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
3	HEM	A	501	1,4	-	0/10/54/54	0/0/8/8
4	KH4	B	101	3,2	-	0/26/33/34	0/1/1/1
3	HEM	C	501	1,4	-	0/10/54/54	0/0/8/8
4	KH4	D	101	3,2	-	0/26/33/34	0/1/1/1
3	HEM	E	501	1,4	-	0/10/54/54	0/0/8/8
4	KH4	F	101	3,2	-	0/26/33/34	0/1/1/1

All (34) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	HEM	C3B-C4B	-7.15	1.45	1.51
3	E	501	HEM	C3B-C4B	-7.06	1.45	1.51
3	C	501	HEM	C3B-C4B	-7.01	1.45	1.51
3	C	501	HEM	C3D-C4D	-4.86	1.45	1.51
3	E	501	HEM	C3D-C4D	-4.86	1.45	1.51
3	A	501	HEM	C3D-C4D	-4.74	1.45	1.51
3	A	501	HEM	C2C-C1C	-3.69	1.45	1.52
3	E	501	HEM	C2C-C1C	-3.62	1.45	1.52
3	C	501	HEM	C2C-C1C	-3.55	1.45	1.52
3	A	501	HEM	C2B-C1B	-2.02	1.45	1.51
3	E	501	HEM	C2B-C1B	-2.01	1.45	1.51
3	A	501	HEM	C4C-NC	2.02	1.38	1.36
3	A	501	HEM	C3C-CAC	2.03	1.55	1.51
3	C	501	HEM	C3C-CAC	2.07	1.55	1.51
3	A	501	HEM	C3B-CAB	2.12	1.55	1.51
3	E	501	HEM	C3B-CAB	2.13	1.55	1.51
3	E	501	HEM	C4C-NC	2.13	1.38	1.36
3	A	501	HEM	C1C-NC	2.13	1.38	1.36
3	C	501	HEM	C1C-NC	2.18	1.38	1.36
3	C	501	HEM	C3B-CAB	2.23	1.55	1.51
3	E	501	HEM	C1C-NC	2.29	1.38	1.36
3	E	501	HEM	FE-NC	2.39	2.05	1.95
3	C	501	HEM	C4C-NC	2.40	1.39	1.36
3	A	501	HEM	FE-NC	2.52	2.05	1.95
3	A	501	HEM	FE-ND	2.64	2.11	1.97
3	C	501	HEM	FE-NC	2.69	2.06	1.95
3	E	501	HEM	FE-ND	2.75	2.12	1.97
3	C	501	HEM	FE-ND	2.87	2.12	1.97
4	D	101	KH4	C16-N2	5.41	1.44	1.33
4	F	101	KH4	C16-N2	5.45	1.45	1.33
4	B	101	KH4	C16-N2	5.45	1.45	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	101	KH4	C4-C5	5.53	1.47	1.37
4	F	101	KH4	C4-C5	5.58	1.47	1.37
4	B	101	KH4	C4-C5	5.69	1.47	1.37

All (49) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	101	KH4	C11-S1-C22	-3.66	94.94	99.59
4	F	101	KH4	C11-S1-C22	-3.51	95.12	99.59
3	E	501	HEM	C3C-CAC-CBC	-2.75	120.23	124.46
3	E	501	HEM	CBA-CAA-C2A	-2.67	107.75	112.53
4	B	101	KH4	C4-C5-C22	-2.63	123.61	129.63
3	C	501	HEM	C3C-CAC-CBC	-2.49	120.63	124.46
3	A	501	HEM	C3C-CAC-CBC	-2.49	120.64	124.46
3	A	501	HEM	CBA-CAA-C2A	-2.49	108.07	112.53
3	C	501	HEM	CBA-CAA-C2A	-2.28	108.44	112.53
3	C	501	HEM	C3B-CAB-CBB	-2.18	121.12	124.46
4	B	101	KH4	C15-C14-C13	-2.17	108.73	112.31
3	A	501	HEM	C3B-CAB-CBB	-2.13	121.19	124.46
3	C	501	HEM	C3B-C4B-NB	-2.11	107.60	111.63
3	E	501	HEM	C3B-CAB-CBB	-2.09	121.25	124.46
4	B	101	KH4	C11-S1-C22	-2.06	96.97	99.59
3	A	501	HEM	C3B-C4B-NB	-2.05	107.70	111.63
3	E	501	HEM	C3B-C4B-NB	-2.03	107.76	111.63
4	F	101	KH4	C4-N3-C2	2.05	108.95	105.71
3	E	501	HEM	C2C-C1C-CHC	2.07	126.83	123.68
3	A	501	HEM	C2C-C1C-CHC	2.11	126.88	123.68
3	A	501	HEM	C2D-C3D-C4D	2.25	105.31	101.50
3	C	501	HEM	C2C-C1C-CHC	2.25	127.10	123.68
3	E	501	HEM	C2D-C3D-C4D	2.26	105.33	101.50
4	D	101	KH4	C4-N3-C2	2.27	109.29	105.71
3	C	501	HEM	C2D-C3D-C4D	2.33	105.45	101.50
4	B	101	KH4	C4-N3-C2	2.42	109.53	105.71
4	D	101	KH4	C5-C22-S1	2.60	118.18	115.20
4	B	101	KH4	C5-C22-S1	2.61	118.19	115.20
4	F	101	KH4	C5-C22-S1	2.65	118.24	115.20
3	C	501	HEM	CMD-C2D-C3D	2.80	126.75	114.35
3	A	501	HEM	CMD-C2D-C3D	2.90	127.16	114.35
3	E	501	HEM	CMD-C2D-C3D	2.90	127.17	114.35
4	B	101	KH4	O2P-C19-C18	2.92	115.24	110.55
3	A	501	HEM	C3B-C4B-CHC	2.93	127.28	123.16
3	E	501	HEM	C3B-C4B-CHC	2.93	127.29	123.16

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	501	HEM	C3B-C4B-CHC	3.00	127.38	123.16
3	C	501	HEM	CMC-C2C-C3C	3.92	126.31	116.53
3	A	501	HEM	CMC-C2C-C3C	3.95	126.39	116.53
3	E	501	HEM	CMC-C2C-C3C	3.97	126.45	116.53
3	C	501	HEM	CMB-C2B-C3B	4.09	126.74	116.53
3	C	501	HEM	CAD-C3D-C4D	4.19	127.24	112.47
3	E	501	HEM	CMB-C2B-C3B	4.19	126.99	116.53
3	A	501	HEM	CAD-C3D-C4D	4.21	127.33	112.47
3	E	501	HEM	CAD-C3D-C4D	4.26	127.51	112.47
3	A	501	HEM	CMB-C2B-C3B	4.28	127.23	116.53
3	E	501	HEM	CAD-C3D-C2D	4.85	127.16	113.22
3	C	501	HEM	CAD-C3D-C2D	4.90	127.31	113.22
3	A	501	HEM	CAD-C3D-C2D	4.92	127.36	113.22
4	F	101	KH4	O2P-C19-C18	12.05	129.91	110.55

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	101	KH4	4	0
4	D	101	KH4	4	0
4	F	101	KH4	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(Å <sup>2</sup> )	Q<0.9
1	A	417/446 (93%)	0.07	8 (1%) 70 70	31, 48, 69, 88	0
1	C	409/446 (91%)	-0.05	4 (0%) 84 85	31, 46, 64, 83	0
1	E	407/446 (91%)	0.11	12 (2%) 55 55	34, 56, 87, 114	0
2	B	75/90 (83%)	0.75	10 (13%) 4 3	48, 70, 98, 111	0
2	D	74/90 (82%)	0.43	4 (5%) 29 28	44, 65, 93, 111	0
2	F	74/90 (82%)	0.72	6 (8%) 15 12	52, 71, 94, 108	0
All	All	1456/1608 (90%)	0.13	44 (3%) 54 54	31, 52, 84, 114	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	8	ILE	4.9
2	F	59	GLU	4.8
1	A	10	ASP	4.5
2	F	79	GLY	4.3
2	B	59	GLU	3.7
2	F	9	ARG	3.5
2	B	77	LEU	3.4
2	F	76	VAL	3.4
1	A	7	PRO	3.4
1	A	117	GLN	3.2
1	E	212	ALA	3.2
1	A	9	PRO	3.1
1	C	212	ALA	3.0
2	F	77	LEU	2.9
1	A	184	ASP	2.9
1	E	413	ALA	2.9
2	B	79	GLY	2.8
1	E	410	ALA	2.7
1	E	117	GLN	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	79	GLY	2.7
1	E	21	ARG	2.7
1	C	21	ARG	2.7
1	E	213	LYS	2.7
1	E	385	GLY	2.7
2	B	75	ALA	2.5
1	E	25	VAL	2.4
1	A	213	LYS	2.4
1	C	12	ILE	2.4
1	E	386	GLU	2.3
1	E	186	ASP	2.3
1	A	25	VAL	2.3
2	B	64	ASP	2.2
2	D	58	ALA	2.2
2	B	66	PHE	2.2
1	E	140	ARG	2.2
1	E	409	GLU	2.1
2	D	76	VAL	2.1
2	B	78	ILE	2.1
1	C	215	THR	2.1
2	F	63	ARG	2.1
2	B	36	PHE	2.0
2	D	67	GLU	2.0
2	B	63	ARG	2.0
2	B	76	VAL	2.0

## 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( $\text{\AA}^2$ )	Q<0.9
4	KH4	F	101	28/29	0.94	0.20	1.60	37,42,67,71	0
3	HEM	A	501	43/43	0.98	0.18	0.35	32,33,34,35	0
3	HEM	E	501	43/43	0.98	0.17	0.32	36,37,38,39	0
4	KH4	D	101	28/29	0.97	0.18	0.32	30,36,53,54	0
3	HEM	C	501	43/43	0.99	0.17	0.06	31,32,33,33	0
4	KH4	B	101	28/29	0.96	0.17	-0.06	33,39,69,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.