



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 03:58 am GMT

PDB ID : 3BK9  
Title : H55A mutant of tryptophan 2,3-dioxygenase from *Xanthomonas campestris*  
Authors : Bruckmann, C.; Mowat, C.G.  
Deposited on : 2007-12-06  
Resolution : 2.15 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 : 1.7.2 (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

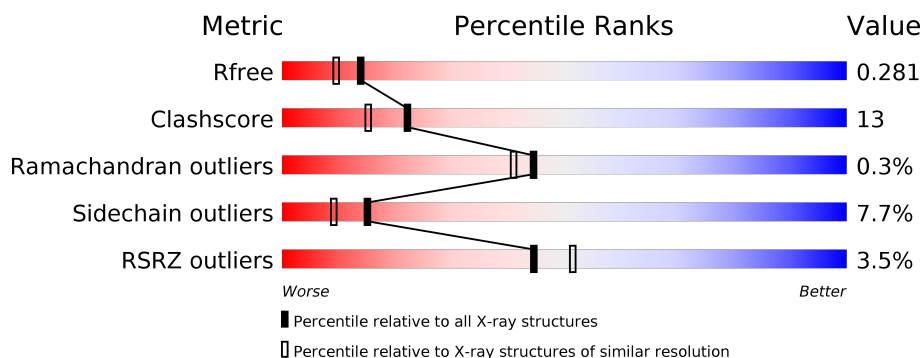
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1170 (2.16-2.16)
Clashscore	112137	1278 (2.16-2.16)
Ramachandran outliers	110173	1256 (2.16-2.16)
Sidechain outliers	110143	1255 (2.16-2.16)
RSRZ outliers	101464	1175 (2.16-2.16)

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	306	 64% 20% 6% • 9%
1	B	306	 60% 25% 6% 8%
1	C	306	 58% 25% 5% 11%
1	D	306	 59% 23% • 14%
1	E	306	 53% 26% 5% • 15%
1	F	306	 55% 23% • • 18%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	306	
1	H	306	

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
2	TRP	D	402	-	-	-	X
2	TRP	E	402	-	-	-	X
2	TRP	G	402	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 19073 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 Tryptophan 2,3-dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	277	Total	C	N	O	S	0	0	0
			2285	1462	404	412	7			
1	B	280	Total	C	N	O	S	0	0	0
			2308	1475	408	418	7			
1	C	272	Total	C	N	O	S	0	0	0
			2244	1437	395	405	7			
1	D	262	Total	C	N	O	S	0	0	0
			2165	1389	381	388	7			
1	E	260	Total	C	N	O	S	0	0	0
			2153	1381	379	386	7			
1	F	250	Total	C	N	O	S	0	0	0
			2071	1330	362	372	7			
1	G	262	Total	C	N	O	S	0	0	0
			2172	1393	384	388	7			
1	H	257	Total	C	N	O	S	0	0	0
			2120	1362	372	379	7			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	55	ALA	HIS	ENGINEERED	UNP Q8PDA8
A	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
A	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
A	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8
A	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
A	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
A	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
A	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
A	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8
B	55	ALA	HIS	ENGINEERED	UNP Q8PDA8
B	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
B	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
B	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8

*Continued on next page...*

*Continued from previous page...*

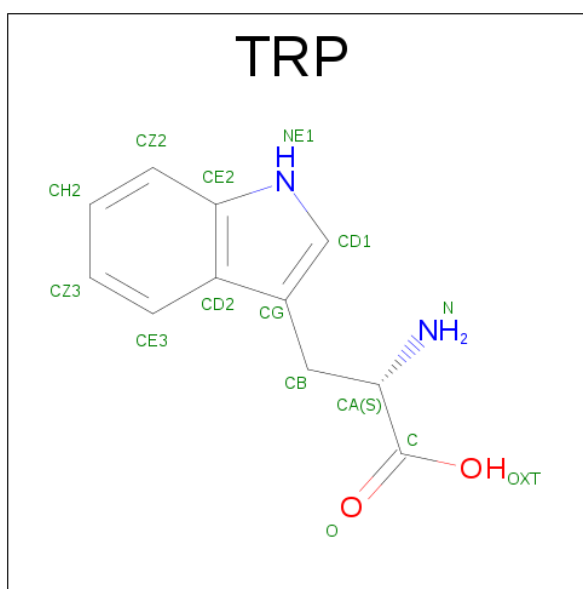
Chain	Residue	Modelled	Actual	Comment	Reference
B	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
B	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
B	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
B	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
B	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8
C	55	ALA	HIS	ENGINEERED	UNP Q8PDA8
C	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
C	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
C	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8
C	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
C	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
C	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
C	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
C	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8
D	55	ALA	HIS	ENGINEERED	UNP Q8PDA8
D	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
D	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
D	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8
D	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
D	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
D	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
D	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
D	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8
E	55	ALA	HIS	ENGINEERED	UNP Q8PDA8
E	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
E	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
E	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8
E	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
E	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
E	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
E	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
E	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8
F	55	ALA	HIS	ENGINEERED	UNP Q8PDA8
F	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
F	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
F	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8
F	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
F	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
F	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
F	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
F	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8
G	55	ALA	HIS	ENGINEERED	UNP Q8PDA8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
G	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
G	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8
G	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
G	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
G	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
G	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
G	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8
H	55	ALA	HIS	ENGINEERED	UNP Q8PDA8
H	299	LEU	-	EXPRESSION TAG	UNP Q8PDA8
H	300	GLU	-	EXPRESSION TAG	UNP Q8PDA8
H	301	HIS	-	EXPRESSION TAG	UNP Q8PDA8
H	302	HIS	-	EXPRESSION TAG	UNP Q8PDA8
H	303	HIS	-	EXPRESSION TAG	UNP Q8PDA8
H	304	HIS	-	EXPRESSION TAG	UNP Q8PDA8
H	305	HIS	-	EXPRESSION TAG	UNP Q8PDA8
H	306	HIS	-	EXPRESSION TAG	UNP Q8PDA8

- Molecule 2 is TRYPTOPHAN (three-letter code: TRP) (formula:  $C_{11}H_{12}N_2O_2$ ).



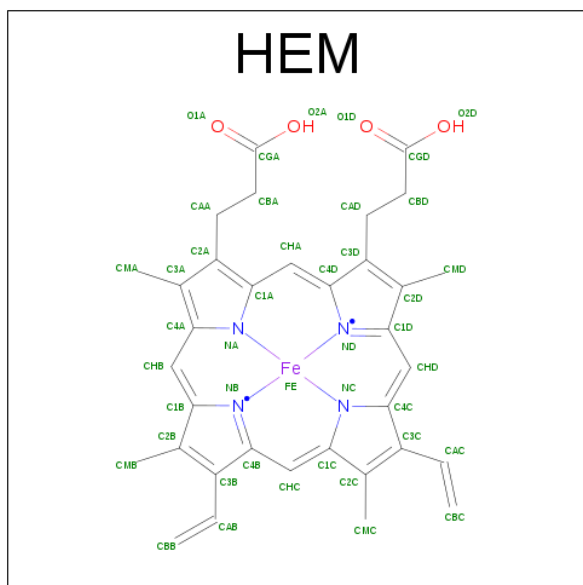
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	11	2	2		
2	B	1	Total	C	N	O	0	0
			15	11	2	2		
2	C	1	Total	C	N	O	0	0
			15	11	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	D	1	Total	C	N	O	0	0
			15	11	2	2		
2	E	1	Total	C	N	O	0	0
			15	11	2	2		
2	F	1	Total	C	N	O	0	0
			15	11	2	2		
2	G	1	Total	C	N	O	0	0
			15	11	2	2		
2	H	1	Total	C	N	O	0	0
			15	11	2	2		
2	A	1	Total	C	N	O	0	0
			15	11	2	2		
2	B	1	Total	C	N	O	0	0
			15	11	2	2		
2	C	1	Total	C	N	O	0	0
			15	11	2	2		
2	D	1	Total	C	N	O	0	0
			15	11	2	2		
2	E	1	Total	C	N	O	0	0
			15	11	2	2		

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



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

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	B	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	D	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
3	F	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	G	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	H	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

- Molecule 4 is water.

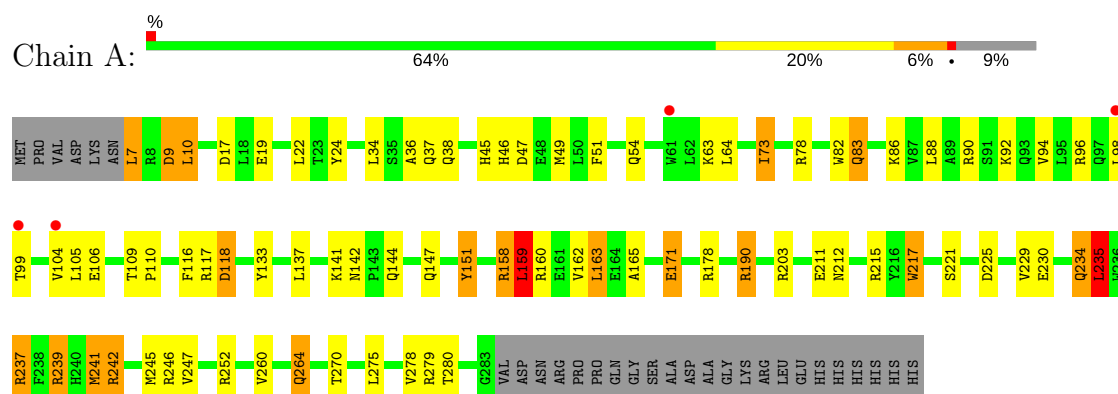
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	170	Total 170	O 170	0	0
4	B	152	Total 152	O 152	0	0
4	C	129	Total 129	O 129	0	0
4	D	124	Total 124	O 124	0	0
4	E	133	Total 133	O 133	0	0
4	F	96	Total 96	O 96	0	0
4	G	113	Total 113	O 113	0	0
4	H	99	Total 99	O 99	0	0



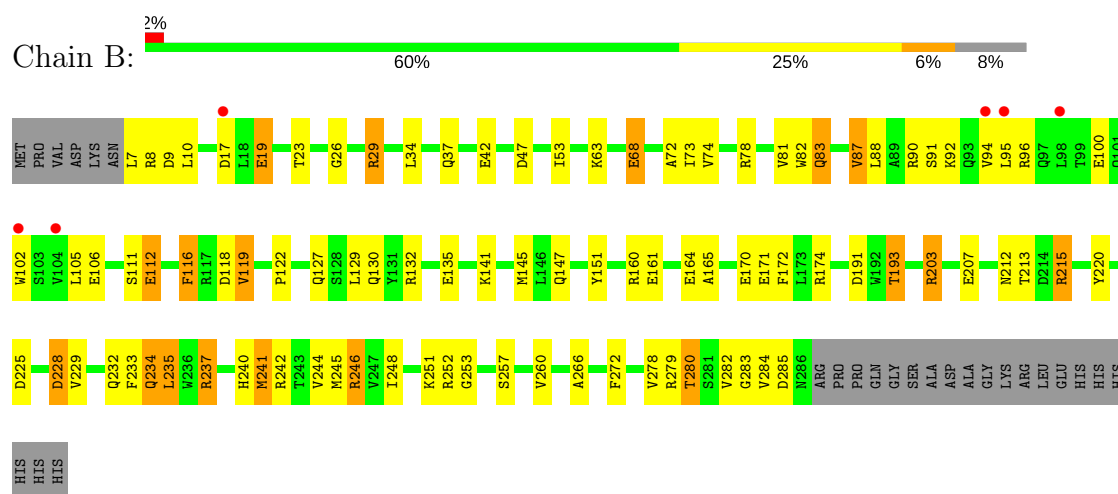
### 3 Residue-property plots

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.

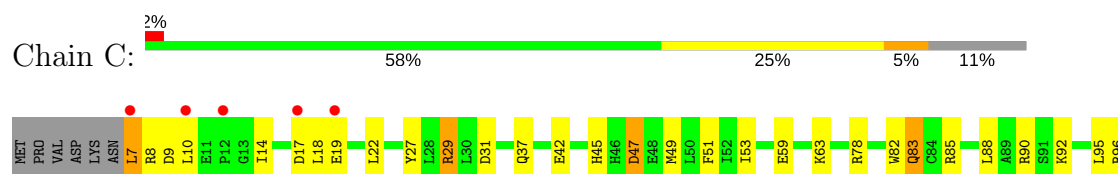
#### • Molecule 1: Tryptophan 2,3-dioxygenase

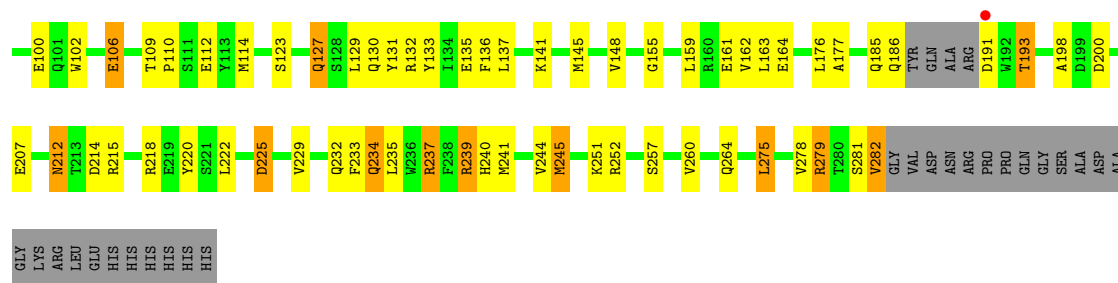


#### • Molecule 1: Tryptophan 2,3-dioxygenase



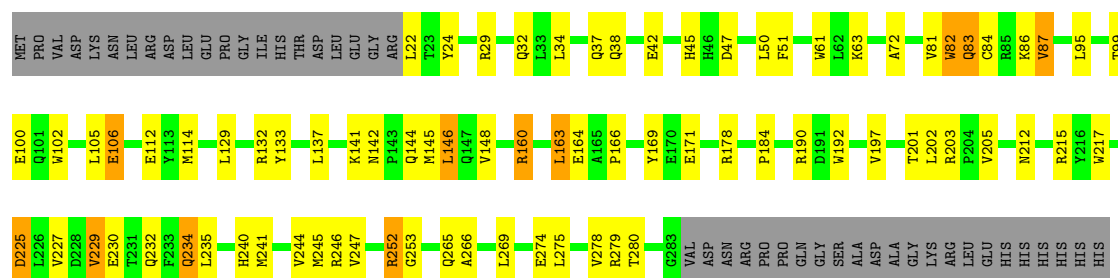
#### • Molecule 1: Tryptophan 2,3-dioxygenase





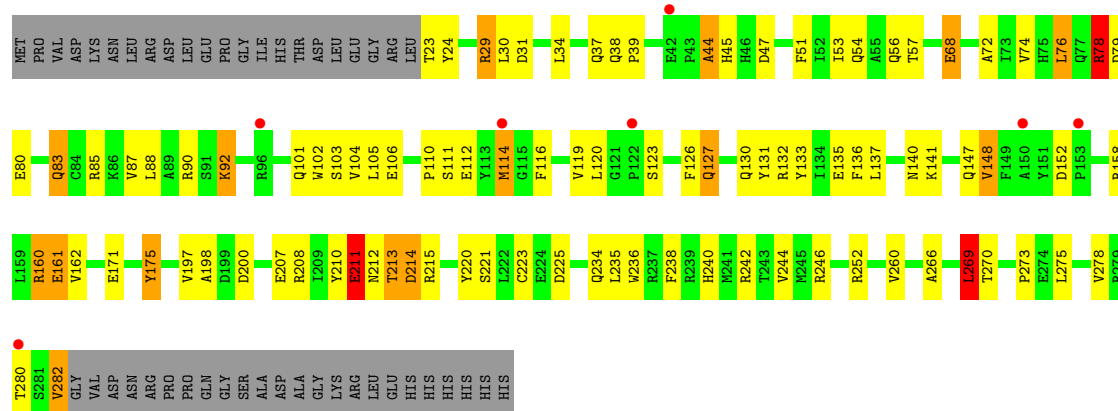
• Molecule 1: Tryptophan 2,3-dioxygenase

Chain D: 59% 23% 14%



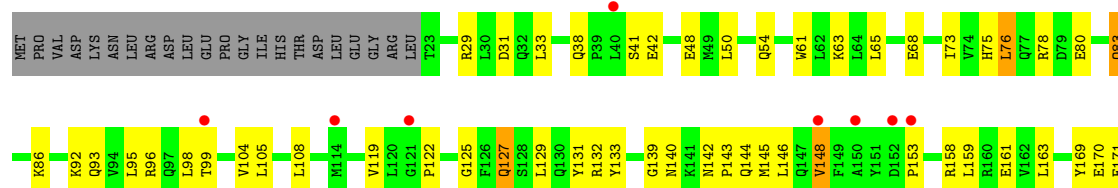
• Molecule 1: Tryptophan 2,3-dioxygenase

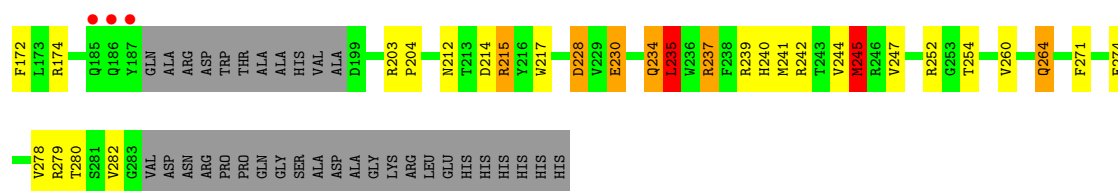
Chain E: 2% 53% 26% 5% 15%



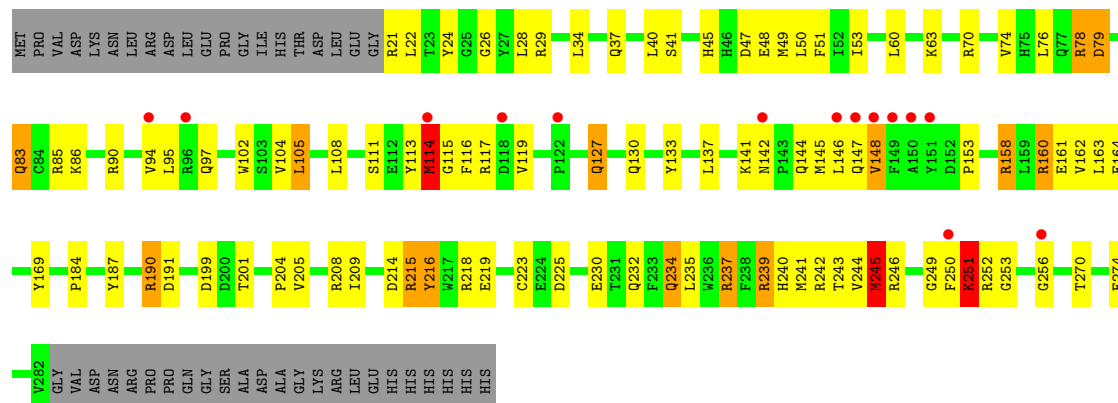
• Molecule 1: Tryptophan 2,3-dioxygenase

Chain F: 4% 55% 23% 18%

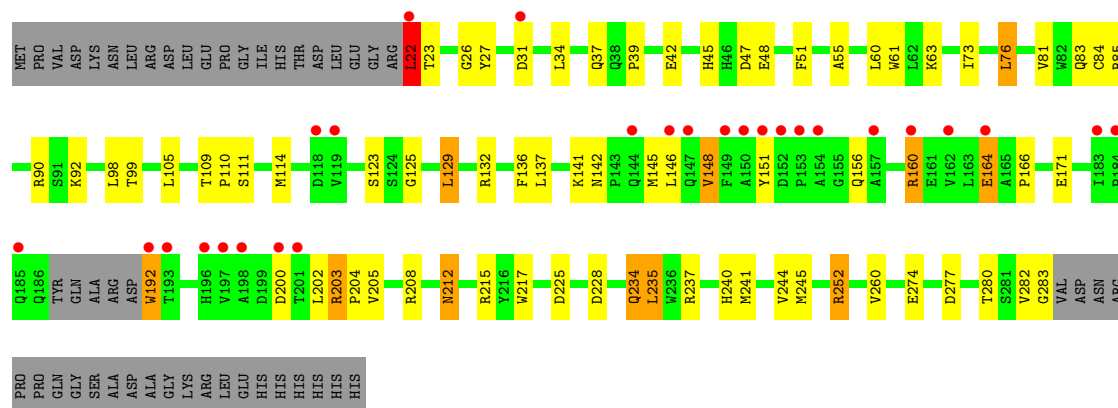




• Molecule 1: Tryptophan 2,3-dioxygenase



• Molecule 1: Tryptophan 2,3-dioxygenase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.22Å 117.61Å 139.28Å 90.00° 95.73° 90.00°	Depositor
Resolution (Å)	54.13 – 2.15 54.13 – 2.15	Depositor EDS
% Data completeness (in resolution range)	96.7 (54.13-2.15) 96.7 (54.13-2.15)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.20 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.200 , 0.285 0.201 , 0.281	Depositor DCC
$R_{free}$ test set	6624 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.6	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 46.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	19073	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.40% 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 ⓘ

### 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	1.46	15/2342 (0.6%)	1.35	26/3174 (0.8%)
1	B	1.52	27/2365 (1.1%)	1.50	24/3206 (0.7%)
1	C	1.35	13/2299 (0.6%)	1.31	19/3115 (0.6%)
1	D	1.41	12/2220 (0.5%)	1.17	9/3009 (0.3%)
1	E	1.36	14/2208 (0.6%)	1.24	14/2993 (0.5%)
1	F	1.28	5/2122 (0.2%)	1.16	13/2872 (0.5%)
1	G	1.27	5/2227 (0.2%)	1.21	14/3018 (0.5%)
1	H	1.18	2/2173 (0.1%)	1.05	5/2944 (0.2%)
All	All	1.36	93/17956 (0.5%)	1.26	124/24331 (0.5%)

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	A	1	0
1	B	0	1
1	G	0	1
All	All	1	2

The worst 5 of 93 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	161	GLU	CG-CD	10.31	1.67	1.51
1	C	106	GLU	CB-CG	8.86	1.69	1.52
1	A	17	ASP	CB-CG	7.96	1.68	1.51
1	D	106	GLU	CB-CG	7.73	1.66	1.52
1	D	106	GLU	CD-OE2	7.31	1.33	1.25

The worst 5 of 124 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	252	ARG	NE-CZ-NH2	-23.35	108.62	120.30
1	B	252	ARG	NE-CZ-NH1	19.05	129.83	120.30
1	B	8	ARG	NE-CZ-NH2	-15.85	112.38	120.30
1	B	8	ARG	NE-CZ-NH1	13.25	126.92	120.30
1	C	279	ARG	NE-CZ-NH2	12.71	126.65	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	17	ASP	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	283	GLY	Peptide
1	G	251	LYS	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	2285	0	2240	70	0
1	B	2308	0	2259	58	0
1	C	2244	0	2201	66	0
1	D	2165	0	2123	64	0
1	E	2153	0	2109	69	0
1	F	2071	0	2033	58	0
1	G	2172	0	2133	77	0
1	H	2120	0	2083	60	0
2	A	30	0	18	3	0
2	B	30	0	18	0	0
2	C	30	0	18	3	0
2	D	30	0	18	0	0
2	E	30	0	18	1	0
2	F	15	0	9	0	0
2	G	15	0	9	1	0
2	H	15	0	9	2	0
3	A	43	0	30	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	43	0	30	4	0
3	C	43	0	30	2	0
3	D	43	0	30	4	0
3	E	43	0	30	7	0
3	F	43	0	30	5	0
3	G	43	0	30	3	0
3	H	43	0	30	5	0
4	A	170	0	0	4	0
4	B	152	0	0	6	0
4	C	129	0	0	7	0
4	D	124	0	0	5	0
4	E	133	0	0	6	0
4	F	96	0	0	2	0
4	G	113	0	0	11	0
4	H	99	0	0	1	0
All	All	19073	0	17538	470	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 13.

The worst 5 of 470 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:114:MET:SD	1:H:283:GLY:HA2	1.77	1.24
1:E:38:GLN:NE2	1:F:29:ARG:HH11	1.37	1.23
1:E:38:GLN:HE22	1:F:29:ARG:NH1	1.38	1.20
1:A:241:MET:CE	1:A:245:MET:HE2	1.75	1.16
1:C:241:MET:SD	1:C:245:MET:HE3	1.86	1.15

There are no symmetry-related clashes.

## 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	275/306 (90%)	269 (98%)	5 (2%)	1 (0%)	38	32
1	B	278/306 (91%)	268 (96%)	9 (3%)	1 (0%)	38	32
1	C	268/306 (88%)	258 (96%)	9 (3%)	1 (0%)	38	32
1	D	260/306 (85%)	250 (96%)	9 (4%)	1 (0%)	38	32
1	E	258/306 (84%)	245 (95%)	13 (5%)	0	100	100
1	F	246/306 (80%)	235 (96%)	9 (4%)	2 (1%)	22	14
1	G	260/306 (85%)	247 (95%)	12 (5%)	1 (0%)	38	32
1	H	253/306 (83%)	241 (95%)	12 (5%)	0	100	100
All	All	2098/2448 (86%)	2013 (96%)	78 (4%)	7 (0%)	44	41

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	284	VAL
1	A	118	ASP
1	G	114	MET
1	C	198	ALA
1	D	184	PRO

### 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	241/266 (91%)	225 (93%)	16 (7%)	19	13
1	B	244/266 (92%)	229 (94%)	15 (6%)	22	16
1	C	238/266 (90%)	220 (92%)	18 (8%)	15	9
1	D	228/266 (86%)	214 (94%)	14 (6%)	22	16
1	E	227/266 (85%)	204 (90%)	23 (10%)	9	4
1	F	220/266 (83%)	204 (93%)	16 (7%)	16	10
1	G	229/266 (86%)	211 (92%)	18 (8%)	14	8
1	H	224/266 (84%)	205 (92%)	19 (8%)	12	7

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1851/2128 (87%)	1712 (92%)	139 (8%)	15 9

5 of 139 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	280	THR
1	E	211	GLU
1	H	137	LEU
1	E	76	LEU
1	E	147	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 74 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	38	GLN
1	E	37	GLN
1	H	45	HIS
1	D	45	HIS
1	D	212	ASN

### 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](#)

21 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	401	1	28,50,50	2.85	12 (42%)	17,82,82	2.97	9 (52%)
2	TRP	A	402	-	11,16,16	1.83	1 (9%)	11,22,22	0.82	0
2	TRP	A	403	-	11,16,16	1.21	2 (18%)	11,22,22	0.83	0
3	HEM	B	401	1,4	28,50,50	3.02	8 (28%)	17,82,82	2.19	4 (23%)
2	TRP	B	402	-	11,16,16	1.70	3 (27%)	11,22,22	1.19	2 (18%)
2	TRP	B	403	-	11,16,16	0.95	1 (9%)	11,22,22	1.06	1 (9%)
3	HEM	C	401	1	28,50,50	2.57	12 (42%)	17,82,82	2.10	3 (17%)
2	TRP	C	402	-	11,16,16	1.61	1 (9%)	11,22,22	0.83	0
2	TRP	C	403	-	11,16,16	1.09	0	11,22,22	0.89	0
3	HEM	D	401	1	28,50,50	2.69	12 (42%)	17,82,82	2.23	6 (35%)
2	TRP	D	402	-	11,16,16	1.81	3 (27%)	11,22,22	0.69	0
2	TRP	D	403	-	11,16,16	1.31	1 (9%)	11,22,22	0.86	0
3	HEM	E	401	1,4	28,50,50	2.20	9 (32%)	17,82,82	2.24	7 (41%)
2	TRP	E	402	-	11,16,16	1.67	3 (27%)	11,22,22	0.88	0
2	TRP	E	403	-	11,16,16	1.29	2 (18%)	11,22,22	0.87	0
3	HEM	F	401	1	28,50,50	2.40	11 (39%)	17,82,82	2.23	8 (47%)
2	TRP	F	402	-	11,16,16	1.35	2 (18%)	11,22,22	0.95	0
3	HEM	G	401	1	28,50,50	2.06	10 (35%)	17,82,82	1.95	5 (29%)
2	TRP	G	402	-	11,16,16	1.44	2 (18%)	11,22,22	0.76	0
3	HEM	H	401	1	28,50,50	2.44	10 (35%)	17,82,82	2.06	7 (41%)
2	TRP	H	402	-	11,16,16	1.35	2 (18%)	11,22,22	0.92	1 (9%)

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	401	1	-	0/6/54/54	0/0/8/8
2	TRP	A	402	-	-	0/3/8/8	0/2/2/2
2	TRP	A	403	-	-	0/3/8/8	0/2/2/2
3	HEM	B	401	1,4	-	0/6/54/54	0/0/8/8
2	TRP	B	402	-	-	0/3/8/8	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TRP	B	403	-	-	0/3/8/8	0/2/2/2
3	HEM	C	401	1	-	0/6/54/54	0/0/8/8
2	TRP	C	402	-	-	0/3/8/8	0/2/2/2
2	TRP	C	403	-	-	0/3/8/8	0/2/2/2
3	HEM	D	401	1	-	0/6/54/54	0/0/8/8
2	TRP	D	402	-	-	0/3/8/8	0/2/2/2
2	TRP	D	403	-	-	0/3/8/8	0/2/2/2
3	HEM	E	401	1,4	-	0/6/54/54	0/0/8/8
2	TRP	E	402	-	-	0/3/8/8	0/2/2/2
2	TRP	E	403	-	-	0/3/8/8	0/2/2/2
3	HEM	F	401	1	-	0/6/54/54	0/0/8/8
2	TRP	F	402	-	-	0/3/8/8	0/2/2/2
3	HEM	G	401	1	-	0/6/54/54	0/0/8/8
2	TRP	G	402	-	-	0/3/8/8	0/2/2/2
3	HEM	H	401	1	-	0/6/54/54	0/0/8/8
2	TRP	H	402	-	-	0/3/8/8	0/2/2/2

The worst 5 of 107 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	401	HEM	C3C-C2C	-9.15	1.28	1.40
3	B	401	HEM	C3B-C2B	-6.29	1.32	1.40
3	C	401	HEM	C3B-C2B	-5.45	1.33	1.40
3	A	401	HEM	C3C-C2C	-5.43	1.33	1.40
3	D	401	HEM	C3C-C2C	-4.60	1.34	1.40

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	401	HEM	CMA-C3A-C4A	-5.70	119.71	128.46
3	A	401	HEM	CBD-CAD-C3D	-5.55	101.88	112.47
3	E	401	HEM	CBD-CAD-C3D	-5.28	102.39	112.47
3	B	401	HEM	CMA-C3A-C4A	-5.11	120.61	128.46
3	A	401	HEM	CMA-C3A-C4A	-5.10	120.62	128.46

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

14 monomers are involved in 40 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	HEM	1	0
2	A	402	TRP	2	0
2	A	403	TRP	1	0
3	B	401	HEM	4	0
3	C	401	HEM	2	0
2	C	402	TRP	3	0
3	D	401	HEM	4	0
3	E	401	HEM	7	0
2	E	403	TRP	1	0
3	F	401	HEM	5	0
3	G	401	HEM	3	0
2	G	402	TRP	1	0
3	H	401	HEM	5	0
2	H	402	TRP	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 [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	277/306 (90%)	-0.07	4 (1%) 75 80	5, 15, 27, 45	0
1	B	280/306 (91%)	-0.03	6 (2%) 64 71	4, 15, 31, 42	0
1	C	272/306 (88%)	0.06	6 (2%) 62 69	10, 19, 41, 54	0
1	D	262/306 (85%)	0.01	0 100 100	9, 19, 38, 44	0
1	E	260/306 (84%)	0.12	7 (2%) 55 63	9, 24, 37, 43	0
1	F	250/306 (81%)	0.33	11 (4%) 35 43	15, 26, 45, 55	0
1	G	262/306 (85%)	0.21	14 (5%) 27 34	9, 25, 42, 47	0
1	H	257/306 (83%)	0.63	27 (10%) 7 10	16, 31, 52, 63	0
All	All	2120/2448 (86%)	0.15	75 (3%) 44 51	4, 22, 41, 63	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	185	GLN	4.6
1	H	119	VAL	4.4
1	C	19	GLU	4.3
1	H	150	ALA	4.3
1	H	192	TRP	4.3

### 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
2	TRP	D	402	15/15	0.92	0.16	4.25	14,18,25,26	0
2	TRP	E	402	15/15	0.84	0.20	2.98	29,31,33,36	0
2	TRP	G	402	15/15	0.84	0.24	2.33	30,35,40,43	0
2	TRP	H	402	15/15	0.82	0.24	1.97	32,36,40,41	0
2	TRP	E	403	15/15	0.86	0.19	1.44	24,27,38,39	0
2	TRP	F	402	15/15	0.94	0.18	1.10	25,27,32,35	0
2	TRP	C	402	15/15	0.92	0.15	0.90	13,21,28,31	0
2	TRP	A	402	15/15	0.94	0.12	0.55	14,16,23,25	0
2	TRP	B	402	15/15	0.96	0.11	0.17	7,9,18,18	0
3	HEM	D	401	43/43	0.97	0.12	0.07	2,9,17,21	0
3	HEM	B	401	43/43	0.98	0.12	0.04	2,5,10,13	0
3	HEM	A	401	43/43	0.98	0.12	0.03	2,8,13,17	0
2	TRP	C	403	15/15	0.94	0.13	0.01	14,18,27,28	0
3	HEM	C	401	43/43	0.97	0.11	-0.20	4,10,16,19	0
2	TRP	A	403	15/15	0.95	0.13	-0.21	14,21,28,28	0
3	HEM	F	401	43/43	0.96	0.12	-0.52	13,20,27,34	0
3	HEM	H	401	43/43	0.96	0.12	-0.52	17,23,27,28	0
3	HEM	E	401	43/43	0.96	0.11	-0.57	11,23,27,34	0
2	TRP	B	403	15/15	0.94	0.11	-0.76	12,20,26,26	0
3	HEM	G	401	43/43	0.95	0.11	-0.83	20,25,31,38	0
2	TRP	D	403	15/15	0.96	0.10	-1.48	11,15,27,30	0

## 6.5 Other polymers

There are no such residues in this entry.