



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 7, 2016 – 05:18 PM EST

PDB ID : 5TIA
Title : Crystal structure of human TDO in complex with Trp, Northeast Structural Genomics Consortium Target HR6161
Authors : Forouhar, F.; Lewis-Ballester, A.; Lew, S.; Karkashon, S.; Seetharaman, J.; Yeh, S.R.; Tong, L.
Deposited on : 2016-10-01
Resolution : 2.44 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

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	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
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)	:	rb-20028442

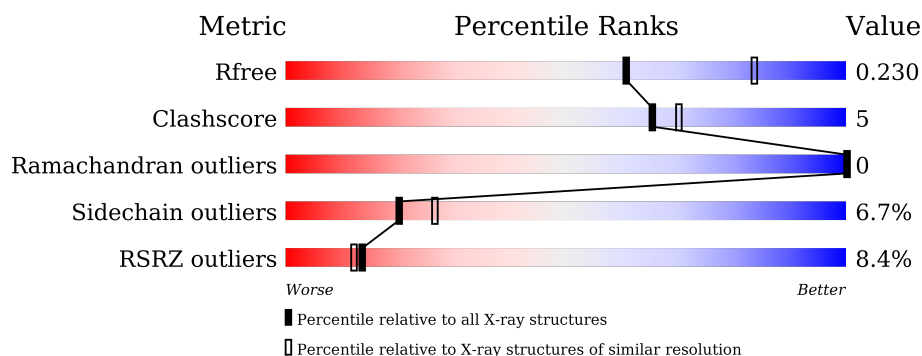
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.44 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	1003 (2.46-2.42)
Clashscore	102246	1071 (2.46-2.42)
Ramachandran outliers	100387	1065 (2.46-2.42)
Sidechain outliers	100360	1065 (2.46-2.42)
RSRZ outliers	91569	1005 (2.46-2.42)

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 ≥ 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	380	<div> <div>6%</div> <div> <div></div> <div>76%</div> <div>11%</div> <div>•</div> <div>11%</div> </div> </div>
1	B	380	<div> <div>5%</div> <div> <div></div> <div>75%</div> <div>11%</div> <div>•</div> <div>12%</div> </div> </div>
1	C	380	<div> <div>12%</div> <div> <div></div> <div>70%</div> <div>12%</div> <div>•</div> <div>15%</div> </div> </div>
1	D	380	<div> <div>7%</div> <div> <div></div> <div>78%</div> <div>12%</div> <div>•</div> <div>8%</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12061 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	337	Total	C	N	O	S	0	0	0
			2837	1822	500	504	11			
1	B	334	Total	C	N	O	S	0	0	0
			2803	1799	486	508	10			
1	C	323	Total	C	N	O	S	0	0	0
			2713	1749	472	482	10			
1	D	349	Total	C	N	O	S	0	0	0
			2942	1889	518	524	11			

There are 32 discrepancies between the modelled and reference sequences:

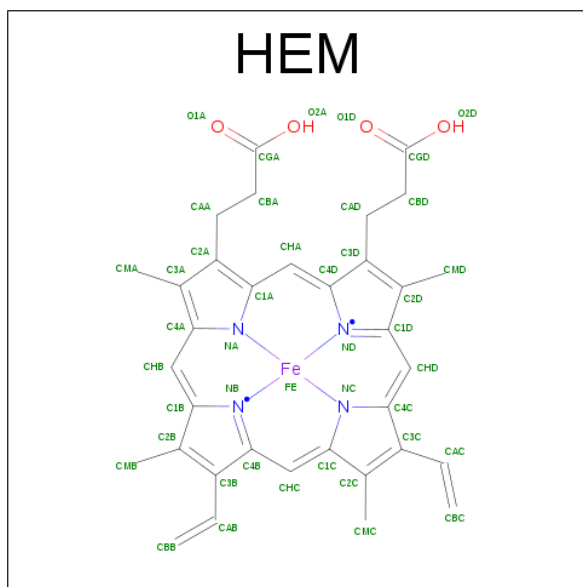
Chain	Residue	Modelled	Actual	Comment	Reference
A	17	MET	-	initiating methionine	UNP P48775
A	390	GLU	-	expression tag	UNP P48775
A	391	HIS	-	expression tag	UNP P48775
A	392	HIS	-	expression tag	UNP P48775
A	393	HIS	-	expression tag	UNP P48775
A	394	HIS	-	expression tag	UNP P48775
A	395	HIS	-	expression tag	UNP P48775
A	396	HIS	-	expression tag	UNP P48775
B	17	MET	-	initiating methionine	UNP P48775
B	390	GLU	-	expression tag	UNP P48775
B	391	HIS	-	expression tag	UNP P48775
B	392	HIS	-	expression tag	UNP P48775
B	393	HIS	-	expression tag	UNP P48775
B	394	HIS	-	expression tag	UNP P48775
B	395	HIS	-	expression tag	UNP P48775
B	396	HIS	-	expression tag	UNP P48775
C	17	MET	-	initiating methionine	UNP P48775
C	390	GLU	-	expression tag	UNP P48775
C	391	HIS	-	expression tag	UNP P48775
C	392	HIS	-	expression tag	UNP P48775
C	393	HIS	-	expression tag	UNP P48775

Continued on next page...

Continued from previous page...

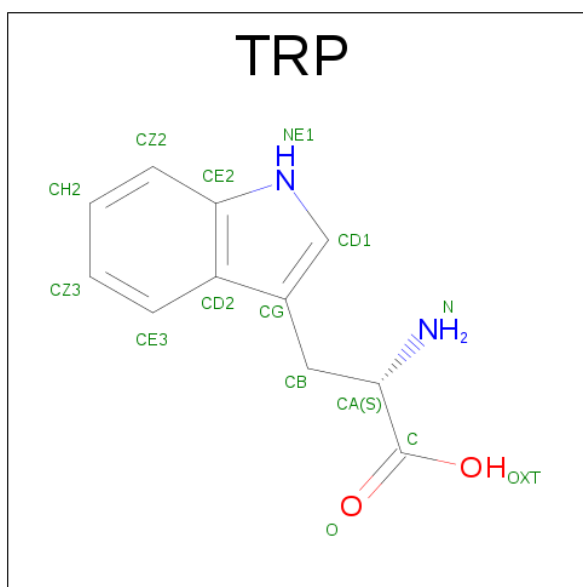
Chain	Residue	Modelled	Actual	Comment	Reference
C	394	HIS	-	expression tag	UNP P48775
C	395	HIS	-	expression tag	UNP P48775
C	396	HIS	-	expression tag	UNP P48775
D	17	MET	-	initiating methionine	UNP P48775
D	390	GLU	-	expression tag	UNP P48775
D	391	HIS	-	expression tag	UNP P48775
D	392	HIS	-	expression tag	UNP P48775
D	393	HIS	-	expression tag	UNP P48775
D	394	HIS	-	expression tag	UNP P48775
D	395	HIS	-	expression tag	UNP P48775
D	396	HIS	-	expression tag	UNP P48775

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

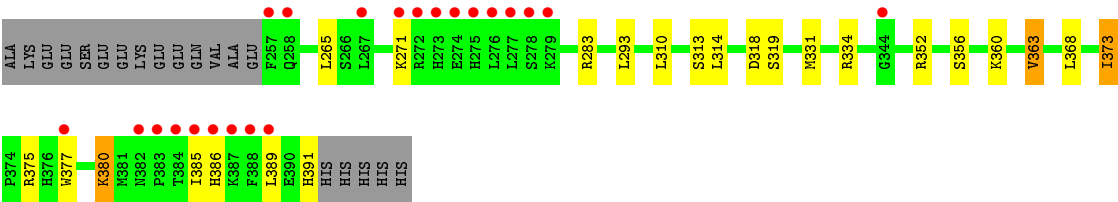
- Molecule 3 is TRYPTOPHAN (three-letter code: TRP) (formula: $C_{11}H_{12}N_2O_2$).



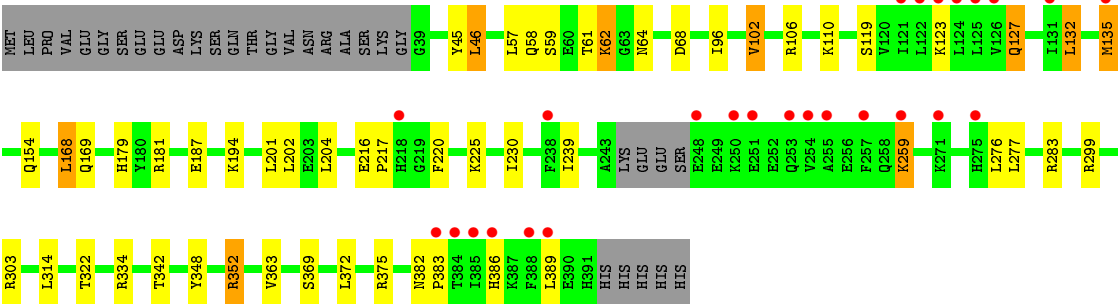
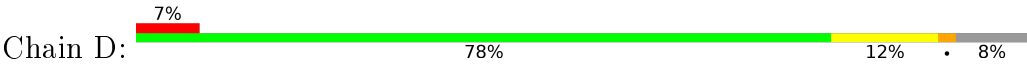
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			15	11	2	2		
3	A	1	Total	C	N	O	0	0
			15	11	2	2		
3	B	1	Total	C	N	O	0	0
			15	11	2	2		
3	B	1	Total	C	N	O	0	0
			15	11	2	2		
3	C	1	Total	C	N	O	0	0
			15	11	2	2		
3	C	1	Total	C	N	O	0	0
			15	11	2	2		
3	D	1	Total	C	N	O	0	0
			15	11	2	2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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	102	Total	O	0	0
			102	102		
4	B	127	Total	O	0	0
			127	127		
4	C	103	Total	O	0	0
			103	103		
4	D	142	Total	O	0	0
			142	142		



● Molecule 1: Tryptophan 2,3-dioxygenase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	144.36 Å 153.56 Å 88.16 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.13 – 2.44 45.92 – 2.44	Depositor EDS
% Data completeness (in resolution range)	99.2 (35.13-2.44) 99.3 (45.92-2.44)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.25 (at 2.45 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.168 , 0.230 0.167 , 0.230	Depositor DCC
R_{free} test set	7363 reflections (10.10%)	DCC
Wilson B-factor (Å ²)	44.8	Xtriage
Anisotropy	0.165	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 60.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	12061	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.26% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 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	0.40	0/2902	0.53	0/3904
1	B	0.41	1/2865 (0.0%)	0.53	0/3853
1	C	0.38	0/2777	0.50	0/3737
1	D	0.40	0/3011	0.55	1/4052 (0.0%)
All	All	0.40	1/11555 (0.0%)	0.53	1/15546 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	251	GLU	CD-OE2	6.51	1.32	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	135	MET	CG-SD-CE	-5.13	91.99	100.20

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	2837	0	2830	22	0
1	B	2803	0	2786	31	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2713	0	2703	32	0
1	D	2942	0	2926	28	0
2	A	43	0	30	2	0
2	B	43	0	30	3	0
2	C	43	0	30	2	0
2	D	43	0	30	3	0
3	A	30	0	18	0	0
3	B	30	0	18	1	0
3	C	30	0	18	0	0
3	D	30	0	18	1	0
4	A	102	0	0	1	0
4	B	127	0	0	3	0
4	C	103	0	0	2	0
4	D	142	0	0	3	0
All	All	12061	0	11437	115	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 5.

The worst 5 of 115 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:150:ALA:HA	1:A:154:GLN:HE22	1.49	0.76
1:D:348:TYR:OH	1:D:352:ARG:NH1	2.25	0.69
1:D:96:ILE:HG22	1:D:102:VAL:HG13	1.75	0.68
1:A:220:PHE:O	1:A:225:LYS:NZ	2.25	0.67
1:A:299:ARG:HH22	1:A:375:ARG:HH22	1.42	0.67

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	333/380 (88%)	324 (97%)	9 (3%)	0	100	100
1	B	330/380 (87%)	324 (98%)	6 (2%)	0	100	100
1	C	317/380 (83%)	310 (98%)	7 (2%)	0	100	100
1	D	345/380 (91%)	337 (98%)	8 (2%)	0	100	100
All	All	1325/1520 (87%)	1295 (98%)	30 (2%)	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	309/348 (89%)	292 (94%)	17 (6%)	27	37
1	B	306/348 (88%)	284 (93%)	22 (7%)	18	23
1	C	297/348 (85%)	274 (92%)	23 (8%)	16	21
1	D	321/348 (92%)	300 (94%)	21 (6%)	21	28
All	All	1233/1392 (89%)	1150 (93%)	83 (7%)	20	26

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	375	ARG
1	C	185	LYS
1	D	276	LEU
1	B	381	MET
1	C	113	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	170	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 ⓘ

12 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	24,50,50	2.34	7 (29%)	16,82,82	1.95	2 (12%)
3	TRP	A	402	-	12,16,16	0.72	0	8,22,22	0.99	0
3	TRP	A	403	-	12,16,16	0.60	0	8,22,22	0.93	0
2	HEM	B	401	1	24,50,50	2.17	5 (20%)	16,82,82	2.02	4 (25%)
3	TRP	B	402	-	12,16,16	0.75	0	8,22,22	1.01	0
3	TRP	B	403	-	12,16,16	0.67	0	8,22,22	0.98	0
2	HEM	C	401	1	24,50,50	2.32	7 (29%)	16,82,82	2.00	4 (25%)
3	TRP	C	402	-	12,16,16	0.70	0	8,22,22	0.97	0
3	TRP	C	403	-	12,16,16	0.68	0	8,22,22	1.02	0
2	HEM	D	401	1	24,50,50	2.30	5 (20%)	16,82,82	2.29	4 (25%)
3	TRP	D	402	-	12,16,16	0.65	0	8,22,22	0.91	0
3	TRP	D	403	-	12,16,16	0.67	0	8,22,22	0.92	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	401	1	-	0/6/54/54	0/0/8/8
3	TRP	A	402	-	-	0/3/8/8	0/2/2/2
3	TRP	A	403	-	-	0/3/8/8	0/2/2/2
2	HEM	B	401	1	-	0/6/54/54	0/0/8/8
3	TRP	B	402	-	-	0/3/8/8	0/2/2/2
3	TRP	B	403	-	-	0/3/8/8	0/2/2/2
2	HEM	C	401	1	-	0/6/54/54	0/0/8/8
3	TRP	C	402	-	-	0/3/8/8	0/2/2/2
3	TRP	C	403	-	-	0/3/8/8	0/2/2/2
2	HEM	D	401	1	-	0/6/54/54	0/0/8/8
3	TRP	D	402	-	-	0/3/8/8	0/2/2/2
3	TRP	D	403	-	-	0/3/8/8	0/2/2/2

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	HEM	C3C-C2C	-5.86	1.32	1.40
2	C	401	HEM	C3C-C2C	-5.24	1.33	1.40
2	D	401	HEM	C3C-C2C	-5.18	1.33	1.40
2	B	401	HEM	C3C-C2C	-4.76	1.34	1.40
2	A	401	HEM	C3B-C2B	-4.23	1.35	1.40

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	401	HEM	CBD-CAD-C3D	-6.78	100.58	112.47
2	D	401	HEM	CBD-CAD-C3D	-6.67	100.77	112.47
2	B	401	HEM	CBD-CAD-C3D	-5.97	101.99	112.47
2	C	401	HEM	CBD-CAD-C3D	-5.40	103.00	112.47
2	D	401	HEM	CBA-CAA-C2A	-3.61	106.14	112.49

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	HEM	2	0
2	B	401	HEM	3	0
3	B	403	TRP	1	0
2	C	401	HEM	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	401	HEM	3	0
3	D	402	TRP	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, 95th 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 ²)	Q<0.9
1	A	337/380 (88%)	0.01	22 (6%) 22 20	32, 48, 88, 129	0
1	B	334/380 (87%)	0.07	18 (5%) 29 28	31, 46, 87, 125	0
1	C	323/380 (85%)	0.48	47 (14%) 3 2	32, 53, 104, 134	0
1	D	349/380 (91%)	0.17	26 (7%) 17 16	32, 47, 95, 158	0
All	All	1343/1520 (88%)	0.18	113 (8%) 14 12	31, 48, 93, 158	0

The worst 5 of 113 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	385	ILE	7.7
1	D	383	PRO	5.6
1	C	384	THR	5.2
1	B	169	GLN	4.9
1	C	385	ILE	4.6

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, 95th 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(\AA^2)	Q<0.9
3	TRP	D	403	15/15	0.98	0.14	0.39	35,41,70,71	0
3	TRP	C	403	15/15	0.95	0.14	0.24	37,51,58,60	0
3	TRP	D	402	15/15	0.98	0.12	0.20	32,39,45,46	0
3	TRP	C	402	15/15	0.96	0.13	-0.19	46,51,58,60	0
2	HEM	D	401	43/43	0.98	0.14	-0.37	29,40,49,51	0
3	TRP	A	403	15/15	0.97	0.11	-0.54	30,45,55,63	0
2	HEM	A	401	43/43	0.97	0.13	-0.55	30,48,61,64	0
3	TRP	B	402	15/15	0.98	0.12	-0.55	32,45,49,51	0
3	TRP	A	402	15/15	0.96	0.13	-0.61	41,51,58,58	0
2	HEM	C	401	43/43	0.96	0.13	-0.72	29,43,57,59	0
2	HEM	B	401	43/43	0.98	0.12	-0.79	30,40,45,47	0
3	TRP	B	403	15/15	0.96	0.11	-1.22	35,41,51,56	0

6.5 Other polymers [i](#)

There are no such residues in this entry.