



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:18 AM GMT

PDB ID : 3ABB
Title : Crystal structure of CYP105D6
Authors : Xu, L.H.; Fushinobu, S.; Takamatsu, S.; Wakagi, T.; Ikeda, H.; Shoun, H.
Deposited on : 2009-12-04
Resolution : 2.30 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
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

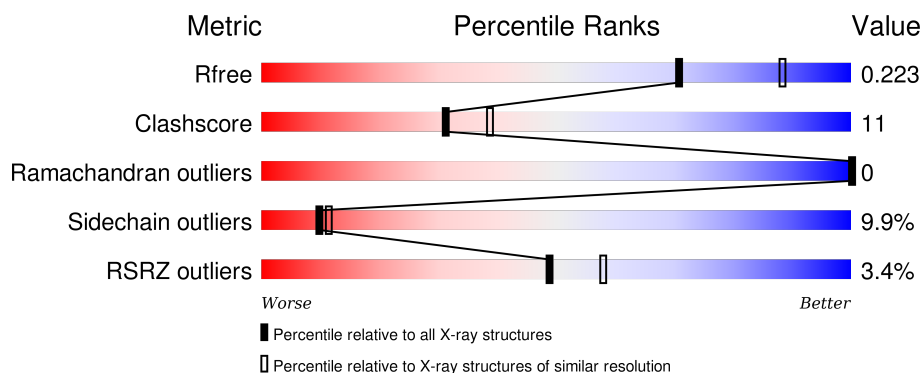
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

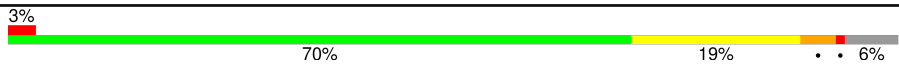
The reported resolution of this entry is 2.30 Å.

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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	408	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3274 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 hydroxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	383	Total	C	N	O	S	0	0	0
			2954	1874	527	543	10			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	405	HIS	-	EXPRESSION TAG	UNP Q79ZT5
A	406	HIS	-	EXPRESSION TAG	UNP Q79ZT5
A	407	HIS	-	EXPRESSION TAG	UNP Q79ZT5
A	408	HIS	-	EXPRESSION TAG	UNP Q79ZT5

- 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		

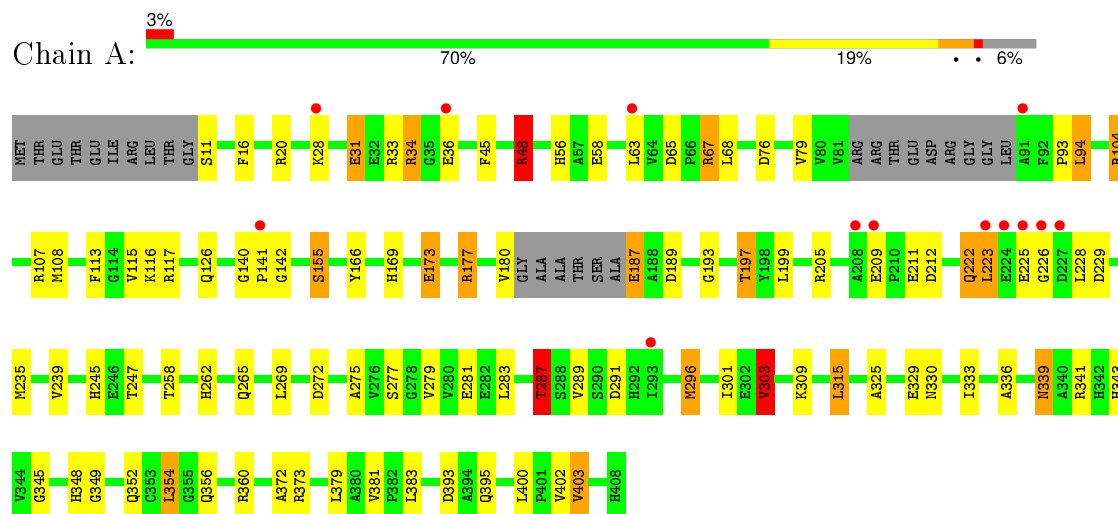
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	277	Total 277	O 277	0	0

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 errors 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 hydroxylase



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	67.53Å 67.53Å 182.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.77 – 2.30 33.20 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.77-2.30) 100.0 (33.20-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.83 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.160 , 0.221 0.164 , 0.223	Depositor DCC
R_{free} test set	1132 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	28.1	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 43.4	EDS
Estimated twinning fraction	0.031 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 22212 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3274	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	1.06	4/3023 (0.1%)	1.05	14/4115 (0.3%)

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	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	31	GLU	CG-CD	6.95	1.62	1.51
1	A	325	ALA	CA-CB	6.19	1.65	1.52
1	A	372	ALA	CA-CB	6.14	1.65	1.52
1	A	58	GLU	CB-CG	5.32	1.62	1.52

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	34	ARG	NE-CZ-NH2	-11.49	114.56	120.30
1	A	48	ARG	NE-CZ-NH2	-9.00	115.80	120.30
1	A	48	ARG	NE-CZ-NH1	8.55	124.58	120.30
1	A	108	MET	CG-SD-CE	-8.04	87.33	100.20
1	A	296	MET	CG-SD-CE	-7.38	88.38	100.20
1	A	34	ARG	NE-CZ-NH1	6.99	123.79	120.30
1	A	67	ARG	NE-CZ-NH2	-5.97	117.31	120.30
1	A	104	ARG	NE-CZ-NH2	-5.92	117.34	120.30
1	A	287	THR	N-CA-CB	-5.66	99.55	110.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	373	ARG	NE-CZ-NH2	-5.54	117.53	120.30
1	A	94	LEU	CB-CG-CD2	-5.49	101.67	111.00
1	A	373	ARG	NE-CZ-NH1	5.33	122.97	120.30
1	A	177	ARG	NE-CZ-NH1	5.15	122.87	120.30
1	A	303	VAL	CB-CA-C	-5.12	101.67	111.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	140	GLY	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	2954	0	2924	64	0
2	A	43	0	30	0	0
3	A	277	0	0	13	0
All	All	3274	0	2954	64	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 11.

All (64) 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:104:ARG:HD3	3:A:592:HOH:O	1.42	1.14
1:A:126:GLN:HG3	3:A:599:HOH:O	1.60	1.02
1:A:356:GLN:HE21	1:A:360:ARG:HH12	1.18	0.91
1:A:193:GLY:O	1:A:197:THR:HG22	1.76	0.85
1:A:265:GLN:HE22	1:A:336:ALA:H	1.24	0.85
1:A:343:HIS:HD2	1:A:345:GLY:H	1.25	0.83
1:A:65:ASP:OD1	1:A:67:ARG:HD3	1.82	0.80
1:A:339:ASN:HD21	1:A:341:ARG:HB2	1.47	0.78

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:222:GLN:NE2	1:A:229:ASP:H	1.84	0.75
1:A:11:SER:HB2	1:A:33:ARG:NH2	2.02	0.75
1:A:225:GLU:H	1:A:226:GLY:HA2	1.52	0.74
1:A:180:VAL:HG21	1:A:245:HIS:ND1	2.02	0.73
1:A:48:ARG:HD3	3:A:425:HOH:O	1.88	0.72
1:A:141:PRO:HB2	1:A:403:VAL:HG12	1.73	0.70
1:A:205:ARG:NH2	3:A:677:HOH:O	2.24	0.70
1:A:117:ARG:HH11	1:A:117:ARG:HG2	1.58	0.69
1:A:265:GLN:NE2	1:A:336:ALA:H	1.91	0.68
1:A:287:THR:HG21	3:A:557:HOH:O	1.93	0.67
1:A:76:ASP:O	1:A:296:MET:HE3	1.94	0.67
1:A:225:GLU:N	1:A:226:GLY:HA2	2.10	0.66
1:A:356:GLN:NE2	1:A:360:ARG:HH12	1.95	0.63
1:A:283:LEU:O	1:A:287:THR:HB	1.99	0.62
1:A:11:SER:HB2	1:A:33:ARG:HH22	1.64	0.61
1:A:56:HIS:HD2	3:A:584:HOH:O	1.83	0.60
1:A:343:HIS:CD2	1:A:345:GLY:H	2.14	0.59
1:A:126:GLN:CG	3:A:599:HOH:O	2.32	0.59
1:A:343:HIS:HD2	1:A:345:GLY:N	1.98	0.59
1:A:209:GLU:HG2	3:A:636:HOH:O	2.04	0.57
1:A:262:HIS:HE1	3:A:418:HOH:O	1.88	0.56
1:A:142:GLY:HA3	1:A:402:VAL:O	2.06	0.56
1:A:356:GLN:HE21	1:A:360:ARG:NH1	1.98	0.55
1:A:247:THR:HG23	3:A:523:HOH:O	2.06	0.55
1:A:113:PHE:CE2	1:A:354:LEU:HD13	2.43	0.53
1:A:301:ILE:HG22	1:A:303:VAL:HG22	1.91	0.53
1:A:281:GLU:OE1	1:A:343:HIS:HE1	1.93	0.52
1:A:155:SER:CB	1:A:173:GLU:HG2	2.41	0.51
1:A:222:GLN:HE21	1:A:229:ASP:H	1.58	0.51
1:A:223:LEU:CD1	1:A:228:LEU:HB2	2.41	0.51
1:A:287:THR:HG23	1:A:289:VAL:HG23	1.92	0.51
1:A:31:GLU:OE1	1:A:34:ARG:HD2	2.12	0.49
1:A:76:ASP:O	1:A:296:MET:CE	2.60	0.48
1:A:339:ASN:ND2	1:A:341:ARG:HB2	2.22	0.48
1:A:20:ARG:HD2	1:A:395:GLN:HE21	1.80	0.47
1:A:126:GLN:HG3	3:A:662:HOH:O	2.13	0.47
1:A:187:GLU:CB	3:A:654:HOH:O	2.63	0.46
1:A:155:SER:HB3	1:A:173:GLU:HG2	1.98	0.45
1:A:258:THR:O	1:A:262:HIS:HD2	2.00	0.45
1:A:11:SER:CB	1:A:33:ARG:HH22	2.27	0.45
1:A:349:GLY:H	1:A:352:GLN:HE21	1.64	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:45:PHE:HE1	1:A:79:VAL:HG23	1.84	0.43
1:A:222:GLN:HE22	1:A:229:ASP:H	1.62	0.43
1:A:63:LEU:HB3	1:A:348:HIS:O	2.18	0.43
1:A:225:GLU:N	1:A:226:GLY:CA	2.80	0.42
1:A:262:HIS:CE1	3:A:418:HOH:O	2.68	0.42
1:A:275:ALA:O	1:A:279:VAL:HG23	2.19	0.42
1:A:16:PHE:CE2	1:A:315:LEU:HD12	2.55	0.42
1:A:93:PRO:O	1:A:94:LEU:HB2	2.20	0.42
1:A:166:TYR:HA	1:A:169:HIS:CD2	2.55	0.41
1:A:68:LEU:HA	1:A:68:LEU:HD23	1.75	0.41
1:A:235:MET:O	1:A:239:VAL:HG23	2.21	0.41
1:A:339:ASN:ND2	1:A:341:ARG:H	2.18	0.41
1:A:187:GLU:OE1	1:A:189:ASP:HB2	2.20	0.41
1:A:211:GLU:HB3	1:A:212:ASP:H	1.65	0.40
1:A:330:ASN:HB3	1:A:333:ILE:HD12	2.03	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	377/408 (92%)	366 (97%)	11 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	312/330 (94%)	281 (90%)	31 (10%)	10	11

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	LYS
1	A	36	GLU
1	A	48	ARG
1	A	107	ARG
1	A	115	VAL
1	A	116	LYS
1	A	155	SER
1	A	173	GLU
1	A	177	ARG
1	A	187	GLU
1	A	197	THR
1	A	199	LEU
1	A	222	GLN
1	A	223	LEU
1	A	269	LEU
1	A	272	ASP
1	A	277	SER
1	A	287	THR
1	A	291	ASP
1	A	303	VAL
1	A	309	LYS
1	A	315	LEU
1	A	329	GLU
1	A	339	ASN
1	A	354	LEU
1	A	379	LEU
1	A	381	VAL
1	A	383	LEU
1	A	393	ASP
1	A	400	LEU
1	A	403	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	56	HIS
1	A	119	ASN
1	A	222	GLN
1	A	262	HIS
1	A	265	GLN
1	A	339	ASN
1	A	343	HIS
1	A	352	GLN
1	A	356	GLN
1	A	357	ASN
1	A	395	GLN
1	A	405	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 ⓘ

1 ligand is 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	1430	1	30,50,50	2.08	8 (26%)	24,82,82	2.39	8 (33%)

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	1430	1	-	0/10/54/54	0/0/8/8

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1430	HEM	C3B-C4B	-5.72	1.46	1.51
2	A	1430	HEM	C2C-C1C	-4.13	1.44	1.52
2	A	1430	HEM	C3D-C4D	-3.51	1.47	1.51
2	A	1430	HEM	C2B-C1B	-2.54	1.43	1.51
2	A	1430	HEM	C2D-C1D	-2.49	1.43	1.51
2	A	1430	HEM	FE-ND	2.52	2.10	1.97
2	A	1430	HEM	C3B-CAB	3.14	1.57	1.51
2	A	1430	HEM	C1C-NC	4.26	1.41	1.36

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1430	HEM	CMA-C3A-C4A	-2.29	124.58	128.36
2	A	1430	HEM	C4B-CHC-C1C	2.12	129.36	125.82
2	A	1430	HEM	CMD-C2D-C3D	2.88	127.10	114.35
2	A	1430	HEM	C3C-CAC-CBC	3.05	129.13	124.46
2	A	1430	HEM	CMC-C2C-C3C	4.21	127.03	116.53
2	A	1430	HEM	CMB-C2B-C3B	4.40	127.51	116.53
2	A	1430	HEM	CAD-C3D-C2D	4.61	126.47	113.22
2	A	1430	HEM	CAD-C3D-C4D	4.88	129.68	112.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	383/408 (93%)	-0.14	13 (3%) 49 58	11, 21, 42, 55	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	208	ALA	4.5
1	A	227	ASP	3.4
1	A	36	GLU	3.3
1	A	209	GLU	3.3
1	A	226	GLY	3.0
1	A	223	LEU	2.7
1	A	91	ALA	2.6
1	A	224	GLU	2.4
1	A	63	LEU	2.3
1	A	293	ILE	2.1
1	A	225	GLU	2.1
1	A	141	PRO	2.1
1	A	28	LYS	2.0

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
2	HEM	A	1430	43/43	0.99	0.19	0.64	12,16,20,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.