



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 7, 2018 – 04:13 am GMT

PDB ID : 1UT0
Title : CRYSTAL STRUCTURE OF CYTOGLOBIN: THE FOURTH GLOBIN TYPE DISCOVERED IN MAN DISPLAYS HEME HEXA-COORDINATION
Authors : De Sanctis, D.; Dewilde, S.; Pesce, A.; Moens, L.; Ascenzi, P.; Hankeln, T.; Burmester, T.; Bolognesi, M.
Deposited on : 2003-12-02
Resolution : 2.10 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

<https://www.wwpdb.org/validation/2017/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 : 1.7.3 (157068), CSD as539be (2018)
Xtrriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : (not set)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

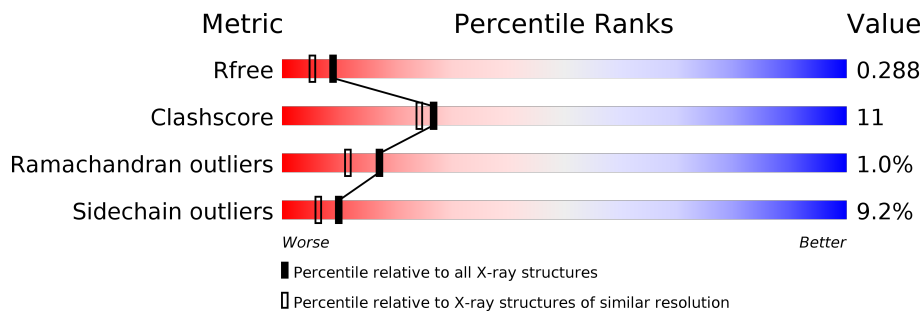
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	111664	4608 (2.10-2.10)
Clashscore	122126	5109 (2.10-2.10)
Ramachandran outliers	120053	5059 (2.10-2.10)
Sidechain outliers	120020	5060 (2.10-2.10)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	190	68% 10% •• 19%
1	B	190	61% 18% •• 19%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2764 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.

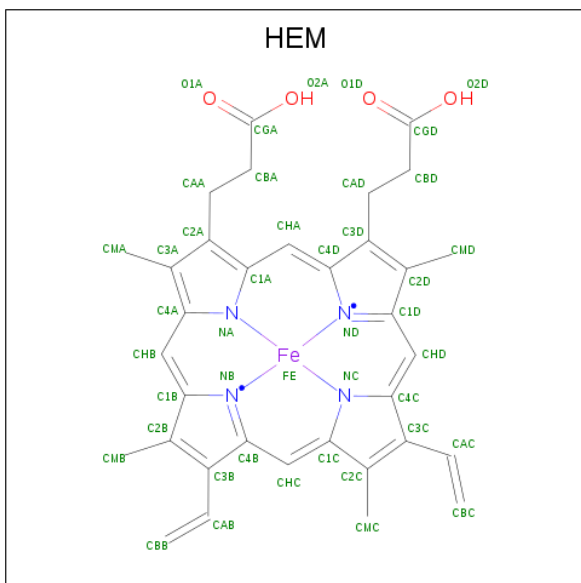
- Molecule 1 is a protein called CYTOGLOBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	154	1232	795	213	220	4	7	0	0
1	B	154	1322	848	235	235	4	0	11	0

There are 4 discrepancies between the modelled and reference sequences:

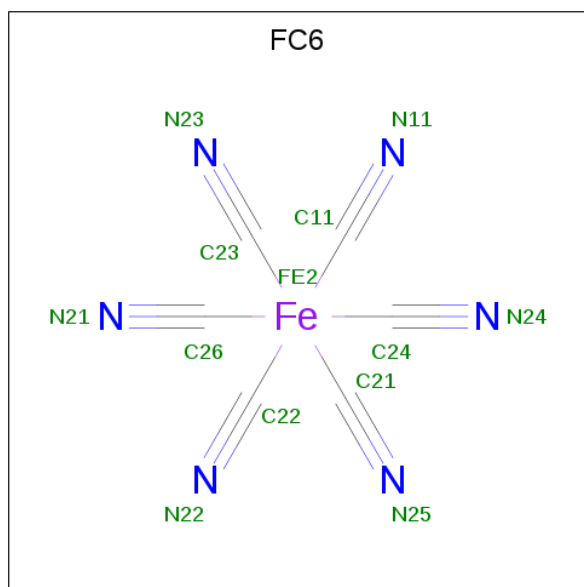
Chain	Residue	Modelled	Actual	Comment	Reference
A	38	SER	CYS	engineered mutation	UNP Q8WWM9
A	83	SER	CYS	engineered mutation	UNP Q8WWM9
B	38	SER	CYS	engineered mutation	UNP Q8WWM9
B	83	SER	CYS	engineered mutation	UNP Q8WWM9

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

- Molecule 3 is HEXACYANOFERRATE(3-) (three-letter code: FC6) (formula: C_6FeN_6).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
3	A	1	Total	C	Fe	N	0	0
			13	6	1	6		
3	B	1	Total	C	Fe	N	0	0
			13	6	1	6		

- Molecule 4 is water.

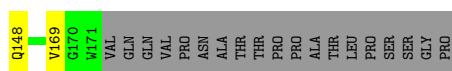
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	57	Total	O	0	0
			57	57		
4	B	41	Total	O	0	0
			41	41		

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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

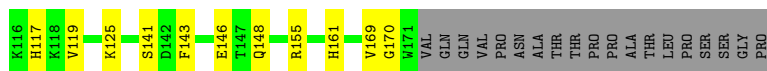
- Molecule 1: CYTOGLOBIN

Chain A: 



- Molecule 1: CYTOGLOBIN

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	48.94Å 70.43Å 100.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.10 50.46 – 2.10	Depositor EDS
% Data completeness (in resolution range)	95.9 (50.00-2.10) 95.9 (50.46-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.94 (at 2.10Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.213 , 0.262 0.265 , 0.288	Depositor DCC
R_{free} test set	1034 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å ²)	36.3	Xtrriage
Anisotropy	0.177	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 40.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	2764	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.97% 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, FC6

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	0.85	2/1263 (0.2%)	0.81	6/1710 (0.4%)
1	B	0.83	1/1355 (0.1%)	0.84	4/1833 (0.2%)
All	All	0.84	3/2618 (0.1%)	0.83	10/3543 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	18	GLU	C-N	21.57	1.83	1.34
1	A	146	GLU	CB-CG	12.32	1.75	1.52
1	A	25	LYS	CG-CD	-12.13	1.11	1.52

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	18	GLU	O-C-N	-16.68	96.02	122.70
1	A	25	LYS	CB-CG-CD	11.76	142.16	111.60
1	B	18	GLU	CA-C-N	8.06	134.93	117.20
1	B	40	ASP	CB-CG-OD2	6.64	124.28	118.30
1	A	98	ASP	CB-CG-OD2	6.54	124.19	118.30
1	A	40	ASP	CB-CG-OD2	6.53	124.18	118.30
1	A	68	ASP	CB-CG-OD2	6.15	123.83	118.30
1	B	98	ASP	CB-CG-OD2	5.17	122.95	118.30
1	A	100	ASP	CB-CG-OD2	5.11	122.90	118.30
1	A	142	ASP	CB-CG-OD2	5.04	122.84	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1232	0	1232	16	0
1	B	1322	0	1322	42	0
2	A	43	0	30	1	0
2	B	43	0	30	4	0
3	A	13	0	0	0	0
3	B	13	0	0	2	0
4	A	57	0	0	6	0
4	B	41	0	0	0	0
All	All	2764	0	2614	58	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 (58) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:18:GLU:C	1:B:19:LEU:N	1.83	1.31
1:B:81[B]:HIS:CD2	1:B:84[B]:ARG:HH12	1.76	1.03
1:B:81[B]:HIS:CD2	1:B:84[B]:ARG:NH1	2.26	1.03
1:B:81[B]:HIS:HD2	1:B:84[B]:ARG:HH12	0.97	0.93
1:B:81[B]:HIS:HA	1:B:84[B]:ARG:HD3	1.51	0.92
1:B:75[B]:SER:OG	1:B:77[B]:GLN:HG2	1.71	0.90
1:B:18:GLU:O	1:B:19:LEU:O	1.89	0.89
1:B:75[B]:SER:OG	1:B:77[B]:GLN:CG	2.42	0.68
1:B:81[B]:HIS:HA	1:B:84[B]:ARG:CD	2.23	0.65
1:B:119:VAL:CG1	2:B:1172:HEM:HAC	2.28	0.63
1:A:46:LEU:HB3	1:A:78:LEU:HD21	1.81	0.62
1:B:18:GLU:O	1:B:19:LEU:N	2.31	0.61
1:B:81[B]:HIS:CA	1:B:84[B]:ARG:HD3	2.29	0.61
1:B:115:LEU:HD21	1:B:169:VAL:HG11	1.83	0.60
1:A:97:HIS:H	1:A:97:HIS:CD2	2.20	0.59
1:B:80[B]:LYS:O	1:B:84[B]:ARG:CD	2.50	0.59
1:B:76[B]:PRO:O	1:B:80[B]:LYS:HG2	2.03	0.58
1:B:18:GLU:O	1:B:19:LEU:CA	2.54	0.55
1:B:81[A]:HIS:O	1:B:85:VAL:HG23	2.07	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:169:VAL:O	1:B:169:VAL:HG12	2.07	0.53
1:B:81[B]:HIS:CG	1:B:84[B]:ARG:NH1	2.74	0.53
1:A:146:GLU:HB3	4:A:2053:HOH:O	2.09	0.52
1:A:37:ASN:ND2	4:A:2010:HOH:O	2.38	0.52
1:A:24:ARG:HD2	1:A:97:HIS:HD2	1.74	0.52
1:A:115:LEU:HD21	1:A:169:VAL:HG11	1.92	0.51
1:B:80[B]:LYS:O	1:B:84[B]:ARG:HD3	2.10	0.51
1:A:140:ALA:C	4:A:2046:HOH:O	2.49	0.51
1:B:18:GLU:C	1:B:19:LEU:CA	2.75	0.50
1:B:18:GLU:O	1:B:19:LEU:C	2.50	0.50
1:B:81[B]:HIS:HA	1:B:84[B]:ARG:CZ	2.41	0.50
1:A:138:GLU:HB3	1:A:139:PHE:CE2	2.47	0.50
1:B:80[B]:LYS:O	1:B:84[B]:ARG:HD2	2.13	0.49
1:B:115:LEU:HD21	1:B:169:VAL:CG1	2.43	0.48
1:B:119:VAL:HG13	2:B:1172:HEM:HAC	1.95	0.48
1:A:146:GLU:CB	4:A:2053:HOH:O	2.61	0.47
1:B:119:VAL:HG11	2:B:1172:HEM:HAC	1.95	0.47
1:B:81[B]:HIS:HA	1:B:84[B]:ARG:NH1	2.29	0.47
1:B:81[B]:HIS:HD2	1:B:84[B]:ARG:NH1	1.70	0.47
1:B:58:GLN:OE1	1:B:64:LYS:NZ	2.48	0.47
1:B:143:PHE:CE2	1:B:148:GLN:HG2	2.48	0.47
1:B:75[B]:SER:HB3	1:B:78[B]:LEU:HB2	1.98	0.46
1:A:117:HIS:HE2	2:A:1172:HEM:CGA	2.28	0.46
1:A:146:GLU:HA	4:A:2053:HOH:O	2.16	0.45
1:B:62:GLN:HE21	1:B:77[A]:GLN:CD	2.20	0.45
1:B:80[B]:LYS:HA	1:B:80[B]:LYS:NZ	2.31	0.45
1:A:140:ALA:HB1	4:A:2046:HOH:O	2.17	0.44
1:A:46:LEU:CB	1:A:78:LEU:HD21	2.48	0.44
1:A:28:GLN:HE22	1:A:97:HIS:CD2	2.35	0.43
1:A:34:LEU:HD22	1:A:139:PHE:CE2	2.54	0.43
1:B:117:HIS:CD2	2:B:1172:HEM:C4D	3.07	0.43
1:B:125:LYS:HE2	3:B:1173:FC6:C26	2.49	0.42
1:B:75[B]:SER:HB3	1:B:78[B]:LEU:CB	2.48	0.42
1:B:155:ARG:HD3	3:B:1173:FC6:N22	2.35	0.41
1:B:30:MET:HE3	1:B:34:LEU:HD13	2.01	0.41
1:B:81[B]:HIS:HA	1:B:84[B]:ARG:NE	2.35	0.41
1:B:50:PHE:CE2	1:B:63:PHE:CE2	3.09	0.41
1:B:81[B]:HIS:CD2	1:B:84[B]:ARG:HH11	2.29	0.41
1:A:143:PHE:CE1	1:A:148:GLN:HG2	2.56	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	152/190 (80%)	148 (97%)	4 (3%)	0	100	100
1	B	163/190 (86%)	156 (96%)	4 (2%)	3 (2%)	9	4
All	All	315/380 (83%)	304 (96%)	8 (2%)	3 (1%)	17	12

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	19	LEU
1	B	141	SER
1	B	170	GLY

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	130/162 (80%)	121 (93%)	9 (7%)	17	13
1	B	140/162 (86%)	123 (88%)	17 (12%)	5	3
All	All	270/324 (83%)	244 (90%)	26 (10%)	10	6

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	LEU
1	A	37	ASN
1	A	80	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	97	HIS
1	A	103	SER
1	A	108	LEU
1	A	138	GLU
1	A	143	PHE
1	A	146	GLU
1	B	18	GLU
1	B	19	LEU
1	B	24	ARG
1	B	46	LEU
1	B	61	SER
1	B	62	GLN
1	B	64	LYS
1	B	75[A]	SER
1	B	75[B]	SER
1	B	79[A]	ARG
1	B	79[B]	ARG
1	B	89	LEU
1	B	97	HIS
1	B	104	SER
1	B	108	LEU
1	B	146	GLU
1	B	161	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	GLN
1	A	37	ASN
1	A	95	ASN
1	A	97	HIS
1	B	62	GLN
1	B	95	ASN
1	B	117	HIS
1	B	148	GLN

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](#)

4 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	1172	1	27,50,50	1.97	5 (18%)	17,82,82	2.08	6 (35%)
3	FC6	A	1173	-	12,12,12	2.49	6 (50%)	0,21,21	0.00	-
2	HEM	B	1172	1	27,50,50	2.19	6 (22%)	17,82,82	1.77	5 (29%)
3	FC6	B	1173	-	12,12,12	2.49	6 (50%)	0,21,21	0.00	-

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	1172	1	-	0/6/54/54	0/0/8/8
3	FC6	A	1173	-	-	0/0/30/30	0/0/0/0
2	HEM	B	1172	1	-	0/6/54/54	0/0/8/8
3	FC6	B	1173	-	-	0/0/30/30	0/0/0/0

All (23) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1172	HEM	C3C-C2C	-4.99	1.33	1.40
2	B	1172	HEM	C3C-C2C	-4.83	1.33	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1172	HEM	C3B-C2B	-4.29	1.34	1.40
2	A	1172	HEM	C3B-C2B	-3.80	1.35	1.40
3	B	1173	FC6	C24-FE2	-3.63	1.82	1.93
3	B	1173	FC6	C22-FE2	-3.57	1.83	1.93
3	A	1173	FC6	C26-FE2	-3.54	1.83	1.93
3	A	1173	FC6	C21-FE2	-3.50	1.83	1.93
3	A	1173	FC6	C11-FE2	-3.47	1.83	1.93
3	B	1173	FC6	C21-FE2	-3.42	1.83	1.93
3	B	1173	FC6	C11-FE2	-3.42	1.83	1.93
3	B	1173	FC6	C23-FE2	-3.41	1.83	1.93
3	B	1173	FC6	C26-FE2	-3.40	1.83	1.93
3	A	1173	FC6	C24-FE2	-3.30	1.83	1.93
3	A	1173	FC6	C23-FE2	-3.25	1.83	1.93
3	A	1173	FC6	C22-FE2	-3.19	1.84	1.93
2	B	1172	HEM	CAA-C2A	2.10	1.55	1.52
2	A	1172	HEM	C3C-CAC	2.72	1.53	1.47
2	A	1172	HEM	C3B-CAB	3.66	1.55	1.47
2	B	1172	HEM	C3B-CAB	3.69	1.55	1.47
2	A	1172	HEM	C3D-C2D	4.33	1.50	1.37
2	B	1172	HEM	C3C-CAC	4.34	1.56	1.47
2	B	1172	HEM	C3D-C2D	4.85	1.52	1.37

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1172	HEM	CBD-CAD-C3D	-5.23	102.50	112.47
2	B	1172	HEM	CBD-CAD-C3D	-4.08	104.69	112.47
2	A	1172	HEM	C1D-C2D-C3D	-3.61	104.49	107.00
2	A	1172	HEM	CMA-C3A-C4A	-2.79	124.18	128.46
2	B	1172	HEM	CMA-C3A-C4A	-2.40	124.78	128.46
2	B	1172	HEM	C1D-C2D-C3D	-2.38	105.34	107.00
2	A	1172	HEM	CAA-CBA-CGA	-2.16	108.96	112.66
2	A	1172	HEM	CMA-C3A-C2A	2.05	128.81	124.94
2	A	1172	HEM	CMC-C2C-C3C	2.18	128.86	124.88
2	B	1172	HEM	C4C-C3C-C2C	2.19	108.43	106.90
2	B	1172	HEM	CMB-C2B-C3B	2.85	130.07	124.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1172	HEM	1	0
2	B	1172	HEM	4	0
3	B	1173	FC6	2	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	18:GLU	C	19:LEU	N	1.83

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.