



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 05:56 pm GMT

PDB ID : 1HBS  
Title : REFINED CRYSTAL STRUCTURE OF DEOXYHEMOGLOBIN S. I. RE-  
STRAINED LEAST-SQUARES REFINEMENT AT 3.0-ANGSTROMS RES-  
OLUTION  
Authors : Padlan, E.A.; Love, W.E.  
Deposited on : 1982-06-02  
Resolution : 3.00 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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

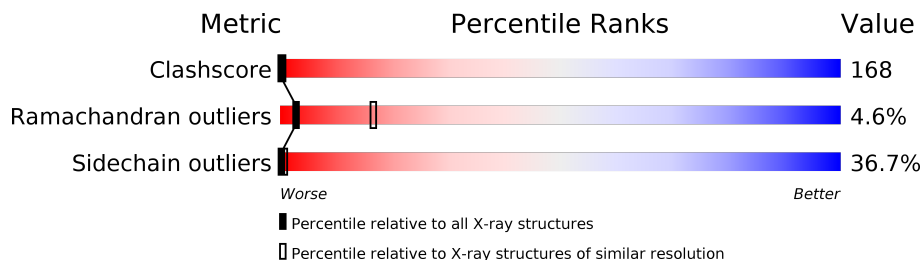
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	141	<div> <div>15%</div> <div>62%</div> <div>21%</div> </div>
1	C	141	<div> <div>9%</div> <div>55%</div> <div>36%</div> </div>
1	E	141	<div> <div>16%</div> <div>52%</div> <div>31%</div> </div>
1	G	141	<div> <div>15%</div> <div>50%</div> <div>35%</div> </div>
2	B	146	<div> <div>12%</div> <div>62%</div> <div>26%</div> </div>
2	D	146	<div> <div>10%</div> <div>66%</div> <div>24%</div> </div>
2	F	146	<div> <div>12%</div> <div>63%</div> <div>24%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	146	

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
3	HEM	A	142	-	-	X	-
3	HEM	D	147	-	-	X	-
3	HEM	E	142	-	-	X	-
3	HEM	G	142	-	-	X	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9104 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 HEMOGLOBIN S (DEOXY) (ALPHA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	141	Total	C	N	O	S	0	0	0
			1069	685	187	194	3			
1	C	141	Total	C	N	O	S	0	0	0
			1069	685	187	194	3			
1	E	141	Total	C	N	O	S	0	0	0
			1069	685	187	194	3			
1	G	141	Total	C	N	O	S	0	0	0
			1069	685	187	194	3			

- Molecule 2 is a protein called HEMOGLOBIN S (DEOXY) (BETA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	146	Total	C	N	O	S	0	0	0
			1121	724	195	199	3			
2	D	146	Total	C	N	O	S	0	0	0
			1121	724	195	199	3			
2	F	146	Total	C	N	O	S	0	0	0
			1121	724	195	199	3			
2	H	146	Total	C	N	O	S	0	0	0
			1121	724	195	199	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	6	VAL	GLU	CONFLICT	UNP P68871
D	6	VAL	GLU	CONFLICT	UNP P68871
F	6	VAL	GLU	CONFLICT	UNP P68871
H	6	VAL	GLU	CONFLICT	UNP P68871

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



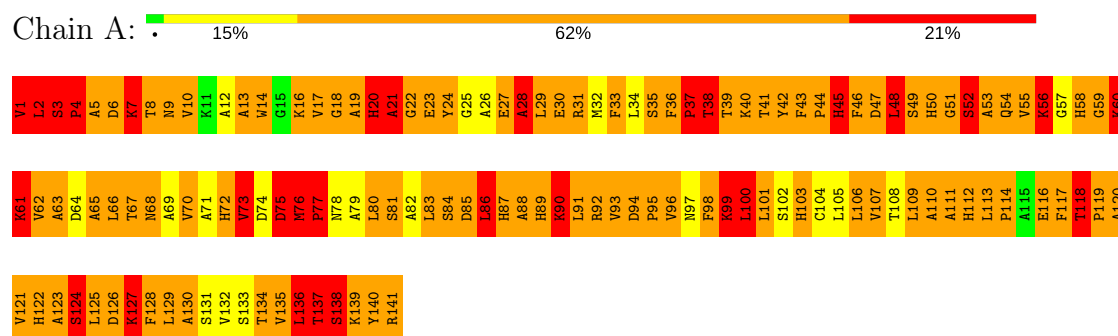
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	E	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	F	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	G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

### 3 Residue-property plots

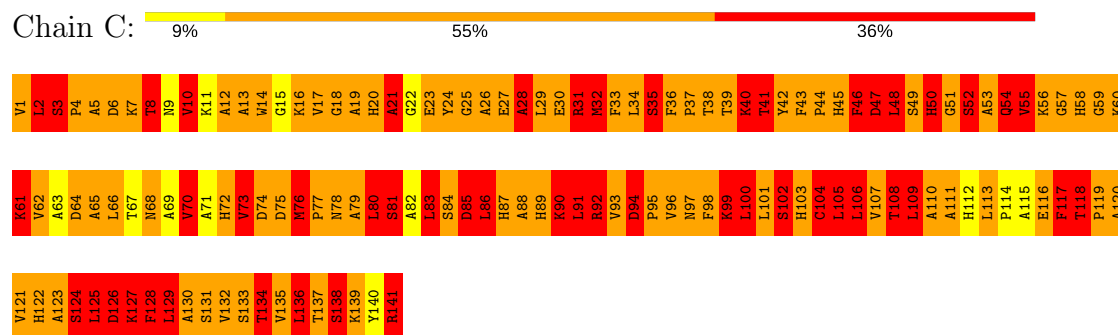
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

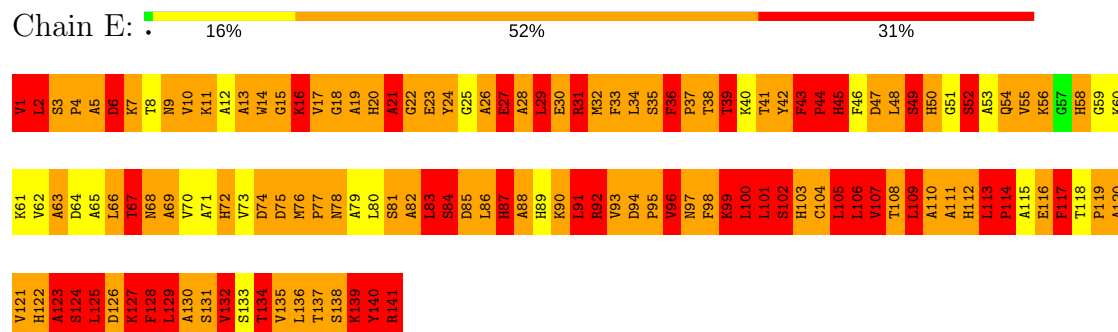
#### • Molecule 1: HEMOGLOBIN S (DEOXY) (ALPHA CHAIN)



#### • Molecule 1: HEMOGLOBIN S (DEOXY) (ALPHA CHAIN)

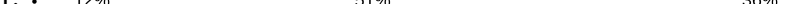


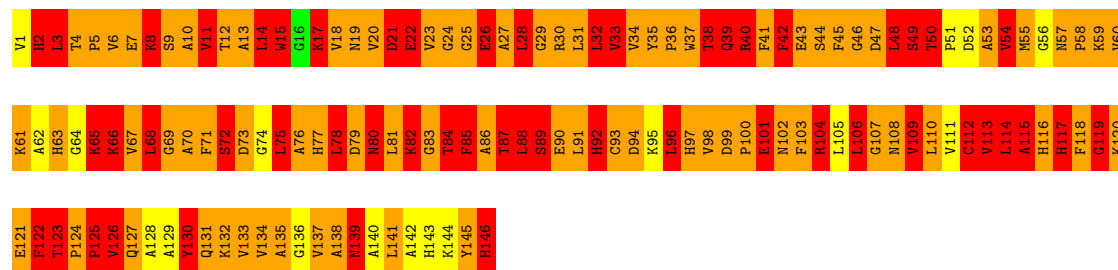
#### • Molecule 1: HEMOGLOBIN S (DEOXY) (ALPHA CHAIN)



E121	K61	V1
F122	A62	H2
T123	H63	L3
P124	G64	T4
P125	K65	P5
V126	K66	V6
Q127	V67	E7
A128	L68	K8
A129	G69	S9
Y130	A70	A10
Q131	F71	V11
K132	S72	T12
V133	D73	A13
V134	G74	L14
A135	L75	V15
G136	A76	G16
V137	H77	K17
A138	L78	V18
M139	D79	M19
A140	N80	V20
L141	L81	D21
A142	K82	E22
H143	G83	V23
K144	T84	G24
Y145	F85	G25
H146	A86	E26
	T87	A27
	L88	L28
	S89	G29
	E90	R30
	L91	L31
	H92	L32
	C93	V33
	D94	V34
	K95	V35
	L96	P36
	H97	W37
	V98	T38
	D99	Q39
	P100	R40
	E101	F41
	N102	F42
	P103	E43
	R104	S44
	L105	F45
	L106	G46
	G107	D47
	N108	L48
	V109	S49
	L110	T50
	V111	P51
	G112	D52
	V113	A53
	L114	V54
	A115	M55
	H116	G56
	H117	N57
	F118	P58
	G119	K59
	V120	W60

- Molecule 2: HEMOGLOBIN S (DEOXY) (BETA CHAIN)

Chain H:  12% 51% 36%





## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.33Å 185.66Å 52.97Å 90.00° 92.69° 90.00°	Depositor
Resolution (Å)	(Not available) – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-3.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	unknown	Depositor
R, $R_{free}$	0.254 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9104	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

## 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	4.15	213/1097 (19.4%)	5.92	486/1491 (32.6%)
1	C	4.30	232/1097 (21.1%)	6.41	521/1491 (34.9%)
1	E	3.93	195/1097 (17.8%)	7.24	524/1491 (35.1%)
1	G	4.08	213/1097 (19.4%)	6.34	533/1491 (35.7%)
2	B	4.12	214/1151 (18.6%)	6.08	512/1564 (32.7%)
2	D	4.35	252/1151 (21.9%)	5.80	499/1564 (31.9%)
2	F	4.23	224/1151 (19.5%)	6.33	516/1564 (33.0%)
2	H	3.98	200/1151 (17.4%)	6.16	548/1564 (35.0%)
All	All	4.14	1743/8992 (19.4%)	6.29	4139/12220 (33.9%)

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	C	0	1
2	F	0	2
All	All	1	3

The worst 5 of 1743 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	23	GLU	CD-OE2	32.40	1.61	1.25
2	D	26	GLU	CD-OE2	25.33	1.53	1.25
2	B	38	THR	C-O	21.93	1.65	1.23
2	D	90	GLU	CD-OE2	21.93	1.49	1.25
2	D	26	GLU	CD-OE1	-21.33	1.02	1.25

The worst 5 of 4139 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	31	ARG	NE-CZ-NH1	103.78	172.19	120.30
1	E	31	ARG	NE-CZ-NH2	-80.58	80.01	120.30
1	G	141	ARG	NE-CZ-NH2	-63.78	88.41	120.30
1	E	42	TYR	CB-CG-CD2	60.94	157.56	121.00
2	F	30	ARG	NE-CZ-NH1	-54.59	93.01	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	20	HIS	CA

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	141	ARG	Sidechain
2	F	23	VAL	Mainchain,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	1069	0	1048	329	0
1	C	1069	0	1035	398	3
1	E	1069	0	1048	334	0
1	G	1069	0	1044	405	0
2	B	1121	0	1100	378	0
2	D	1121	0	1091	467	3
2	F	1121	0	1094	354	3
2	H	1121	0	1091	438	0
3	A	43	0	30	23	0
3	B	43	0	30	14	0
3	C	43	0	30	15	0
3	D	43	0	30	25	0
3	E	43	0	30	24	0
3	F	43	0	30	16	1
3	G	43	0	30	21	0
3	H	43	0	30	7	0
All	All	9104	0	8791	3010	5

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 168.

The worst 5 of 3010 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:76:MET:CA	1:A:76:MET:CB	1.76	1.64
2:D:104:ARG:CG	2:D:104:ARG:CB	1.74	1.64
2:F:29:GLY:HA3	2:F:55:MET:SD	1.32	1.64
2:H:14:LEU:CG	2:H:14:LEU:CD2	1.76	1.63
1:G:32:MET:CA	1:G:32:MET:CB	1.75	1.62

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:53:ALA:CB	2:F:43:GLU:O[1_454]	1.28	0.92
2:D:6:VAL:CG2	2:F:73:ASP:OD2[1_455]	1.92	0.28
1:C:54:GLN:OE1	2:F:46:GLY:CA[1_454]	1.94	0.26
1:C:85:ASP:OD2	2:D:83:GLY:O[1_554]	2.00	0.20
2:D:9:SER:OG	3:F:147:HEM:O2A[1_455]	2.08	0.12

## 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	139/141 (99%)	121 (87%)	15 (11%)	3 (2%)	8	36
1	C	139/141 (99%)	116 (84%)	13 (9%)	10 (7%)	1	6
1	E	139/141 (99%)	113 (81%)	20 (14%)	6 (4%)	3	18
1	G	139/141 (99%)	105 (76%)	24 (17%)	10 (7%)	1	6
2	B	144/146 (99%)	120 (83%)	23 (16%)	1 (1%)	25	67
2	D	144/146 (99%)	120 (83%)	17 (12%)	7 (5%)	2	15

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	144/146 (99%)	123 (85%)	14 (10%)	7 (5%)	2	15
2	H	144/146 (99%)	114 (79%)	22 (15%)	8 (6%)	2	12
All	All	1132/1148 (99%)	932 (82%)	148 (13%)	52 (5%)	3	16

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	18	GLY
1	C	73	VAL
2	D	11	VAL
2	D	40	ARG
2	D	54	VAL

### 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	113/113 (100%)	84 (74%)	29 (26%)	0	3
1	C	113/113 (100%)	66 (58%)	47 (42%)	0	0
1	E	113/113 (100%)	69 (61%)	44 (39%)	0	0
1	G	113/113 (100%)	67 (59%)	46 (41%)	0	0
2	B	118/118 (100%)	73 (62%)	45 (38%)	0	1
2	D	118/118 (100%)	82 (70%)	36 (30%)	0	2
2	F	118/118 (100%)	81 (69%)	37 (31%)	0	1
2	H	118/118 (100%)	63 (53%)	55 (47%)	0	0
All	All	924/924 (100%)	585 (63%)	339 (37%)	0	1

5 of 339 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	125	PRO
1	E	119	PRO

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	80	ASN
1	E	1	VAL
1	E	52	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	80	ASN
1	E	97	ASN
2	H	92	HIS
1	E	58	HIS
1	E	103	HIS

### 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](#)

8 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	142	1	28,50,50	4.39	13 (46%)	17,82,82	6.27	13 (76%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	HEM	B	147	2	28,50,50	5.01	20 (71%)	17,82,82	3.37	11 (64%)
3	HEM	C	142	1	28,50,50	2.94	12 (42%)	17,82,82	5.49	13 (76%)
3	HEM	D	147	2	28,50,50	3.00	10 (35%)	17,82,82	4.11	10 (58%)
3	HEM	E	142	1	28,50,50	4.03	17 (60%)	17,82,82	5.90	13 (76%)
3	HEM	F	147	2	28,50,50	3.77	16 (57%)	17,82,82	4.46	10 (58%)
3	HEM	G	142	-	28,50,50	4.21	15 (53%)	17,82,82	5.02	13 (76%)
3	HEM	H	147	-	28,50,50	3.47	14 (50%)	17,82,82	5.35	11 (64%)

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	142	1	-	0/6/54/54	0/0/8/8
3	HEM	B	147	2	-	0/6/54/54	0/0/8/8
3	HEM	C	142	1	-	0/6/54/54	0/0/8/8
3	HEM	D	147	2	-	0/6/54/54	0/0/8/8
3	HEM	E	142	1	-	0/6/54/54	0/0/8/8
3	HEM	F	147	2	-	0/6/54/54	0/0/8/8
3	HEM	G	142	-	-	0/6/54/54	0/0/8/8
3	HEM	H	147	-	-	0/6/54/54	0/0/8/8

The worst 5 of 117 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	147	HEM	C3C-C2C	-16.28	1.18	1.40
3	A	142	HEM	C3C-C2C	-9.67	1.27	1.40
3	E	142	HEM	C3C-C2C	-8.73	1.28	1.40
3	F	147	HEM	C3C-C2C	-8.50	1.29	1.40
3	F	147	HEM	C3B-C2B	-8.43	1.29	1.40

The worst 5 of 94 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	142	HEM	CAD-CBD-CGD	-15.15	86.77	112.66
3	H	147	HEM	CAA-CBA-CGA	-12.14	91.92	112.66
3	C	142	HEM	CMA-C3A-C2A	-10.19	105.72	124.94
3	G	142	HEM	C4A-C3A-C2A	-9.26	100.56	107.00
3	E	142	HEM	CAD-CBD-CGD	-8.71	97.78	112.66

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 146 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	142	HEM	23	0
3	B	147	HEM	14	0
3	C	142	HEM	15	0
3	D	147	HEM	25	0
3	E	142	HEM	24	0
3	F	147	HEM	16	1
3	G	142	HEM	21	0
3	H	147	HEM	7	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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.