wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 09:30 PM EDT

PDB ID	:	1HAC
Title	:	CROSSLINKED HAEMOGLOBIN
Authors	:	Schumacher, M.A.; Dixon, M.M.; Kluger, R.; Jones, R.T.; Brennan, R.G.
Deposited on	:	1996-03-13
Resolution	:	2.60 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

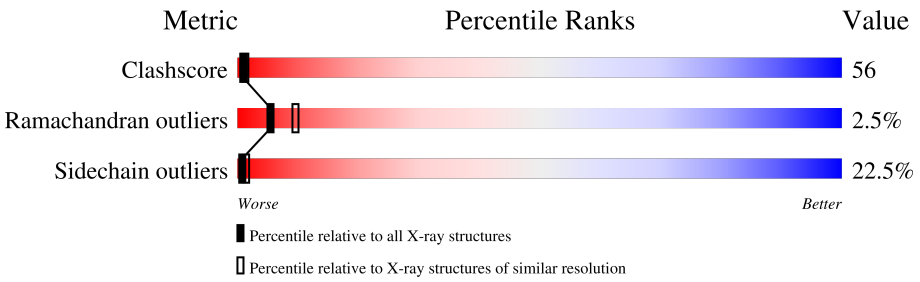
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	141	35% 51% 13% •
1	C	141	35% 46% 17% •
2	B	146	33% 47% 19% •
2	D	146	40% 47% 12% •

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4623 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.

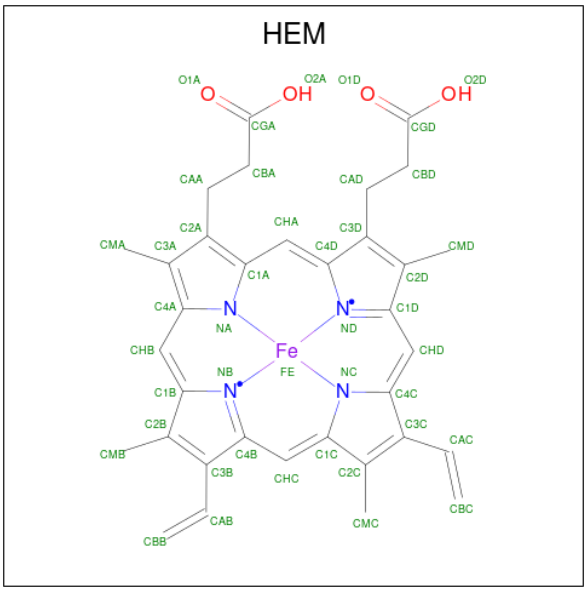
- Molecule 1 is a protein called HEMOGLOBIN A.

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

- Molecule 2 is a protein called HEMOGLOBIN A.

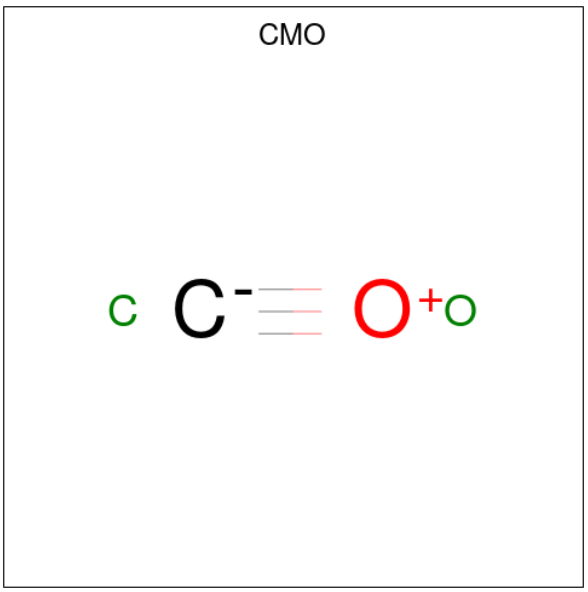
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	146	1123	724	195	201	3	0	0	0
2	D	146	1123	724	195	201	3	0	0	0

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



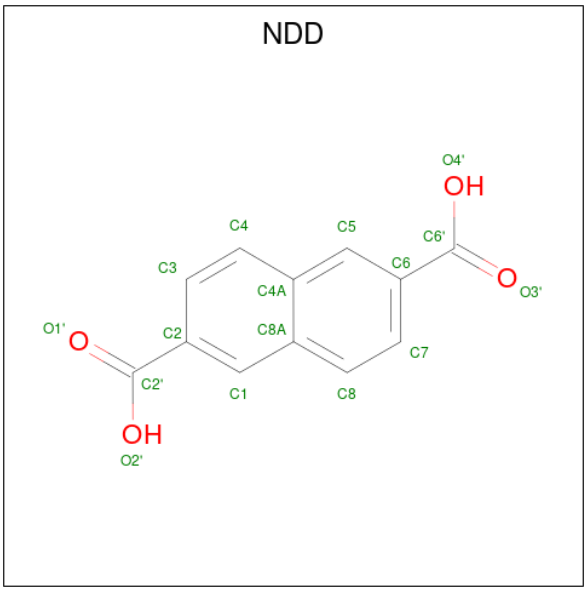
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
3	A	1	43	34	1	4	4	0	0
3	B	1	43	34	1	4	4	0	0
3	C	1	43	34	1	4	4	0	0
3	D	1	43	34	1	4	4	0	0

- Molecule 4 is CARBON MONOXIDE (three-letter code: CMO) (formula: CO).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	2	1	1	0	0
4	B	1	2	1	1	0	0
4	C	1	2	1	1	0	0
4	D	1	2	1	1	0	0

- Molecule 5 is 2,6-DICARBOXYNAPHTHALENE (three-letter code: NDD) (formula: $C_{12}H_8O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	B	1	14	12	2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	7	7	7	0	0
			Total	O		
6	B	7	7	7	0	0
			Total	O		
6	C	14	14	14	0	0
			Total	O		
6	D	17	17	17	0	0

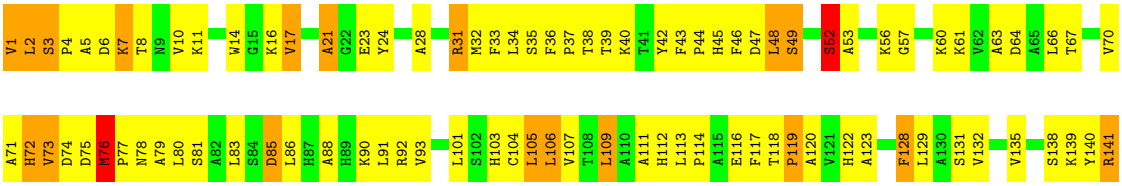
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

Note EDS was not executed.

• Molecule 1: HEMOGLOBIN A

Chain A: 35% 51% 13%



• Molecule 1: HEMOGLOBIN A

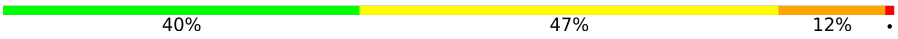
Chain C: 35% 46% 17%

• Molecule 2: HEMOGLOBIN A

Chain B: 33% 47% 19%

• Molecule 2: HEMOGLOBIN A

Chain D:


40%
47%
12%

V1 H2 L3 T4 P5 E6 E7 K8 S9 A10 V11 T12 A13 L14 W15 G16 K17 V18 D21 E22 V23 G24 G25 E26 A27 L28 G29 R30 V33 V34 Y35 P36 W37 T38 Q39 R40 F41 F45 G46 D47 L48 S49 T50 P51 D52 A53 V54 M55 G56 N57 P58 K59 A62 H63 G64 K65
K66 V67 L68 L75 A76 H77 L78 D79 N80 L81 L88 S89 E90 L91 H92 C93 D94 K95 D99 P100 E101 N102 F103 R104 L105 N108 V109 L110 V113 L114 A115 H116 H117 F118 G119 K120 E121 F122 T123 P124 P125 V126 Q127 A128 A129 Y130 Q131 K132 V133 V134 A135 G136 V137 A140
L141 A142 H143 K144 Y145 H146

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.84Å 87.11Å 97.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.60	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
Refinement program	TNT	Depositor
R, R_{free}	0.154 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4623	wwPDB-VP
Average B, all atoms ($Å^2$)	34.0	wwPDB-VP

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: NDD, HEM, CMO

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	1.09	1/1097 (0.1%)	1.42	11/1491 (0.7%)
1	C	1.00	1/1097 (0.1%)	1.31	5/1491 (0.3%)
2	B	1.01	1/1153 (0.1%)	1.29	7/1566 (0.4%)
2	D	1.14	2/1153 (0.2%)	1.32	9/1566 (0.6%)
All	All	1.06	5/4500 (0.1%)	1.34	32/6114 (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	0	1
1	C	0	1
2	B	0	2
2	D	1	2
All	All	1	6

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	143	HIS	CA-C	10.15	1.79	1.52
2	D	143	HIS	N-CA	-9.65	1.27	1.46
1	A	119	PRO	C-N	-7.18	1.17	1.34
1	C	5	ALA	CA-CB	-5.66	1.40	1.52
2	B	111	VAL	CB-CG2	-5.33	1.41	1.52

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	119	PRO	O-C-N	-17.42	94.83	122.70
1	C	36	PHE	C-N-CD	-13.55	90.80	120.60
1	A	119	PRO	CA-C-N	11.47	142.43	117.20
1	C	113	LEU	C-N-CD	-11.21	95.93	120.60
2	B	145	TYR	N-CA-C	-9.16	86.28	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	D	143	HIS	CA

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	119	PRO	Mainchain
2	B	110	LEU	Mainchain
2	B	24	GLY	Mainchain
1	C	117	PHE	Mainchain
2	D	105	LEU	Mainchain

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	1069	0	1072	116	0
1	C	1069	0	1073	139	0
2	B	1123	0	1116	152	0
2	D	1123	0	1115	111	0
3	A	43	0	30	5	0
3	B	43	0	30	6	0
3	C	43	0	30	5	0
3	D	43	0	30	11	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	1	0
5	B	14	0	6	4	0
6	A	7	0	0	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	7	0	0	9	0
6	C	14	0	0	2	1
6	D	17	0	0	7	0
All	All	4623	0	4502	508	1

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 56.

The worst 5 of 508 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)
2:D:143:HIS:C	2:D:143:HIS:CA	1.79	1.48
2:D:4:THR:HG23	2:D:7:GLU:CD	1.69	1.12
1:C:86:LEU:HD12	1:C:90:LYS:HG2	1.21	1.08
2:B:3:LEU:HD12	2:B:7:GLU:HB3	1.35	1.07
1:C:46:PHE:HB3	1:C:48:LEU:HD13	1.37	1.05

All (1) 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)
6:A:202:HOH:O	6:C:203:HOH:O[2_655]	1.12	1.08

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%)	16 (12%)	2 (1%)	11 22
1	C	139/141 (99%)	122 (88%)	13 (9%)	4 (3%)	4 7
2	B	144/146 (99%)	123 (85%)	15 (10%)	6 (4%)	3 3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	144/146 (99%)	125 (87%)	17 (12%)	2 (1%)	11	22
All	All	566/574 (99%)	491 (87%)	61 (11%)	14 (2%)	5	9

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	SER
2	B	12	THR
2	B	2	HIS
2	B	13	ALA
2	B	86	ALA

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	113/113 (100%)	90 (80%)	23 (20%)	1	2
1	C	113/113 (100%)	86 (76%)	27 (24%)	0	1
2	B	118/118 (100%)	86 (73%)	32 (27%)	0	1
2	D	118/118 (100%)	96 (81%)	22 (19%)	1	2
All	All	462/462 (100%)	358 (78%)	104 (22%)	1	1

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	3	SER
1	C	85	ASP
2	D	95	LYS
1	C	34	LEU
1	C	49	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	97	ASN
2	D	77	HIS
2	D	108	ASN
2	D	102	ASN
1	C	9	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

9 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	$\#\|Z\| > 2$
4	CMO	D	201	3	0,1,1	-	-	-		
4	CMO	A	201	3	0,1,1	-	-	-		
3	HEM	C	200	4,1	41,50,50	1.44	4 (9%)	45,82,82	1.94	15 (33%)
4	CMO	B	201	3	0,1,1	-	-	-		
5	NDD	B	400	2	15,15,17	1.09	0	20,20,24	1.03	1 (5%)
3	HEM	D	200	2,4	41,50,50	1.43	3 (7%)	45,82,82	1.79	10 (22%)
3	HEM	A	200	4,1	41,50,50	1.49	6 (14%)	45,82,82	1.86	11 (24%)
3	HEM	B	200	2,4	41,50,50	1.54	6 (14%)	45,82,82	1.86	10 (22%)
4	CMO	C	201	3	0,1,1	-	-	-		

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	C	200	4,1	-	8/12/54/54	-
5	NDD	B	400	2	-	0/4/4/8	0/2/2/2
3	HEM	D	200	2,4	-	7/12/54/54	-
3	HEM	A	200	4,1	-	1/12/54/54	-
3	HEM	B	200	2,4	-	3/12/54/54	-

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	200	HEM	C3C-C2C	-4.62	1.34	1.40
3	B	200	HEM	C3C-C2C	-4.55	1.34	1.40
3	C	200	HEM	C3C-C2C	-4.35	1.34	1.40
3	D	200	HEM	C3C-C2C	-4.25	1.34	1.40
3	A	200	HEM	C3C-CAC	3.55	1.55	1.47

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	200	HEM	C4B-CHC-C1C	5.18	129.39	122.56
3	A	200	HEM	C4C-CHD-C1D	5.16	129.37	122.56
3	B	200	HEM	CBA-CAA-C2A	5.16	121.42	112.62
3	D	200	HEM	CMA-C3A-C4A	-4.59	121.41	128.46
3	D	200	HEM	CBA-CAA-C2A	4.55	120.38	112.62

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	200	HEM	C2D-C3D-CAD-CBD
3	B	200	HEM	C4D-C3D-CAD-CBD
3	C	200	HEM	C1A-C2A-CAA-CBA
3	C	200	HEM	C3A-C2A-CAA-CBA
3	C	200	HEM	C2B-C3B-CAB-CBB

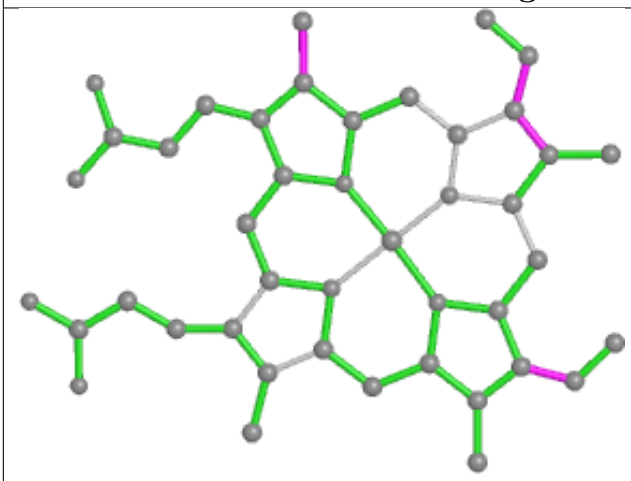
There are no ring outliers.

6 monomers are involved in 32 short contacts:

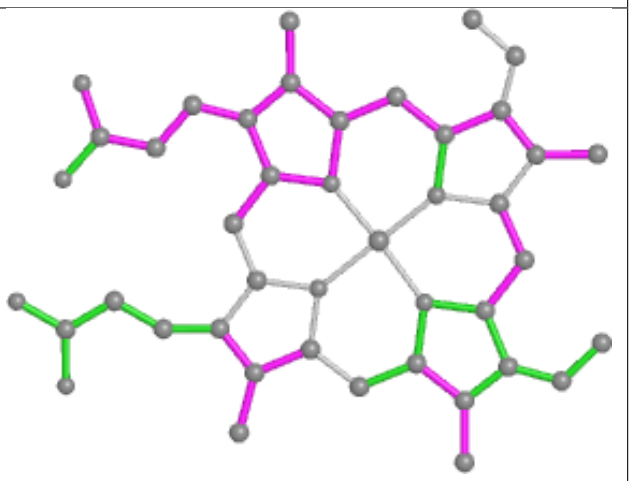
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	201	CMO	1	0
3	C	200	HEM	5	0
5	B	400	NDD	4	0
3	D	200	HEM	11	0
3	A	200	HEM	5	0
3	B	200	HEM	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

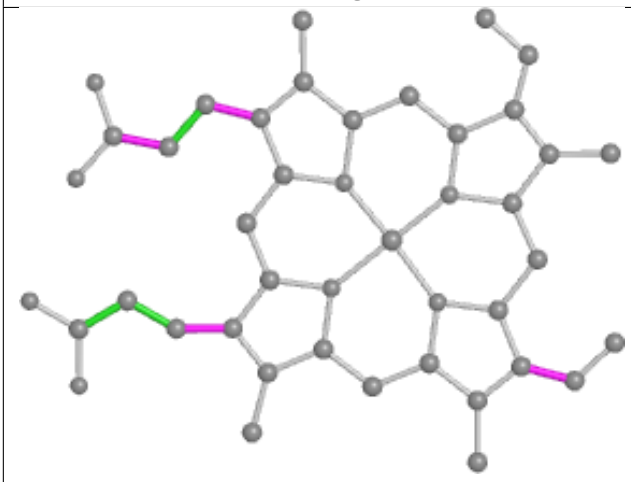
Ligand HEM C 200



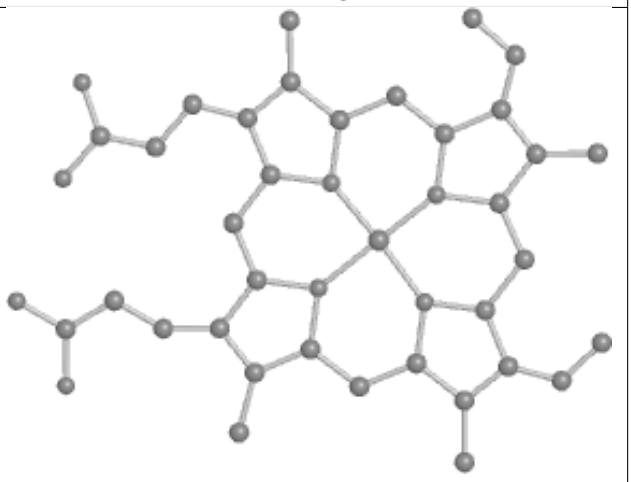
Bond lengths



Bond angles

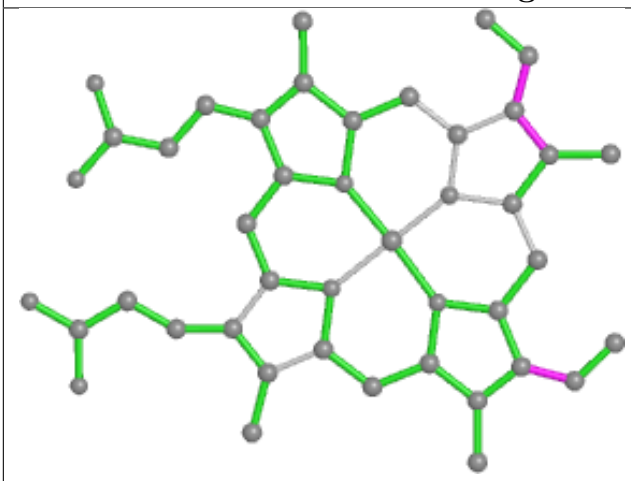


Torsions

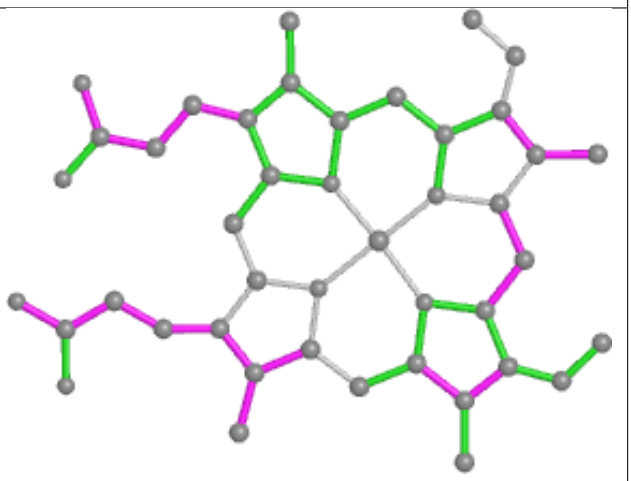


Rings

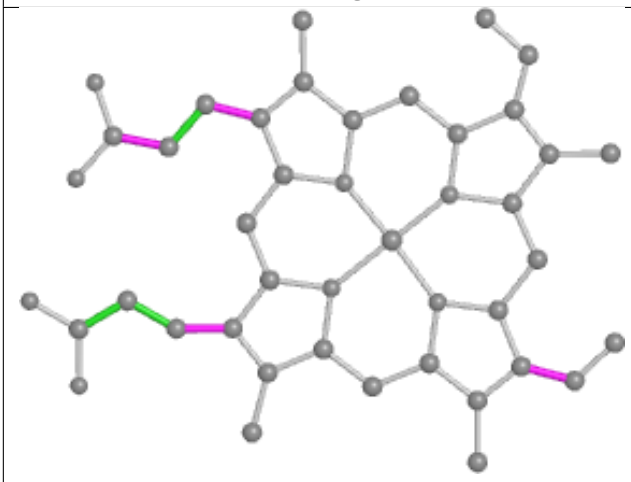
Ligand HEM D 200



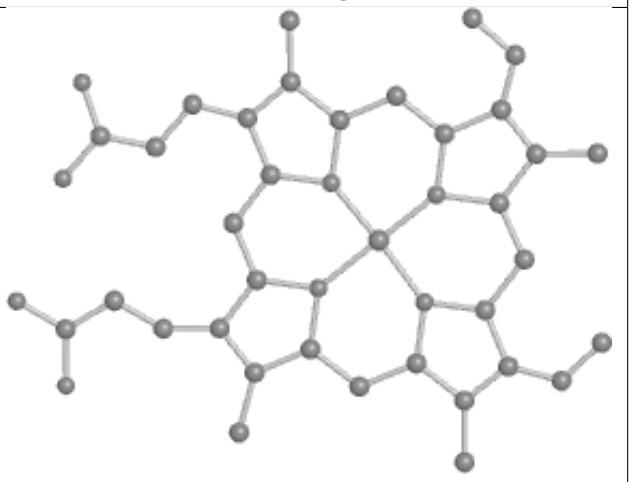
Bond lengths



Bond angles

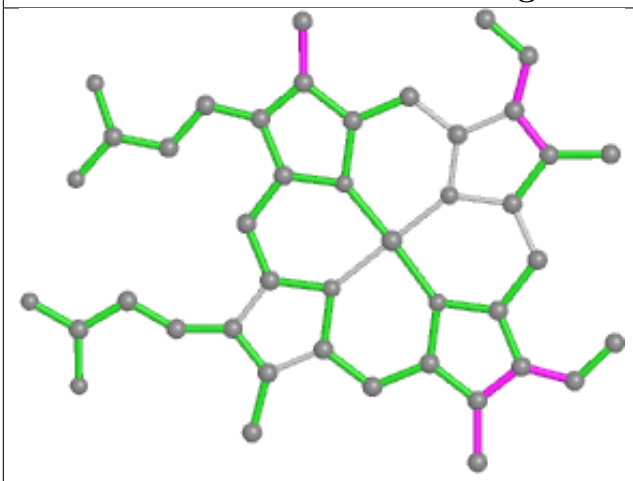


Torsions

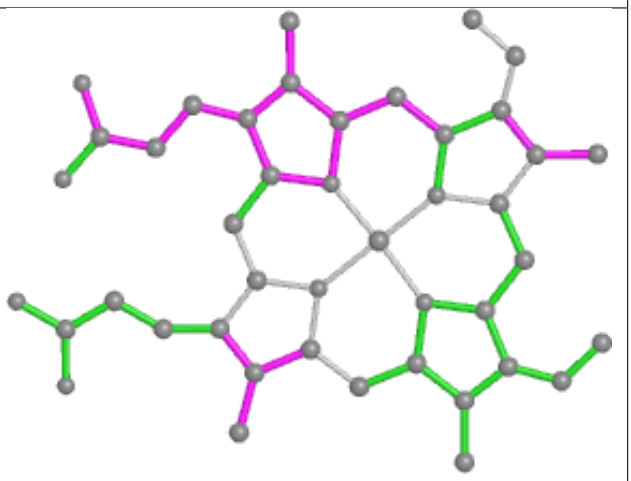


Rings

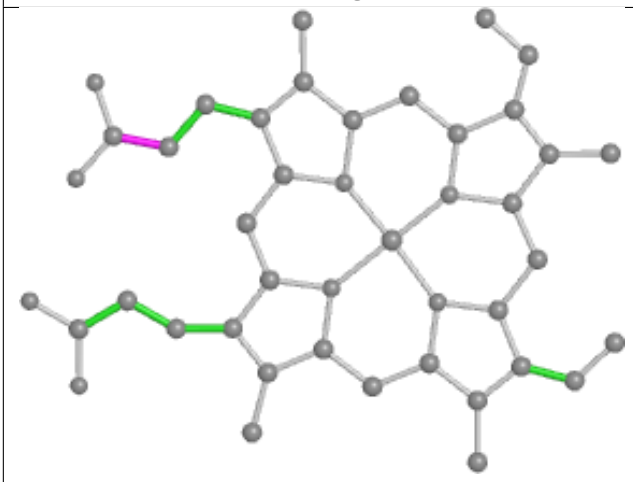
Ligand HEM A 200



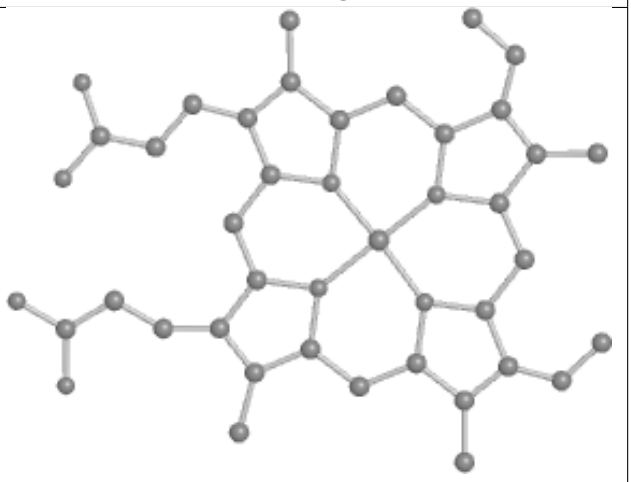
Bond lengths



Bond angles

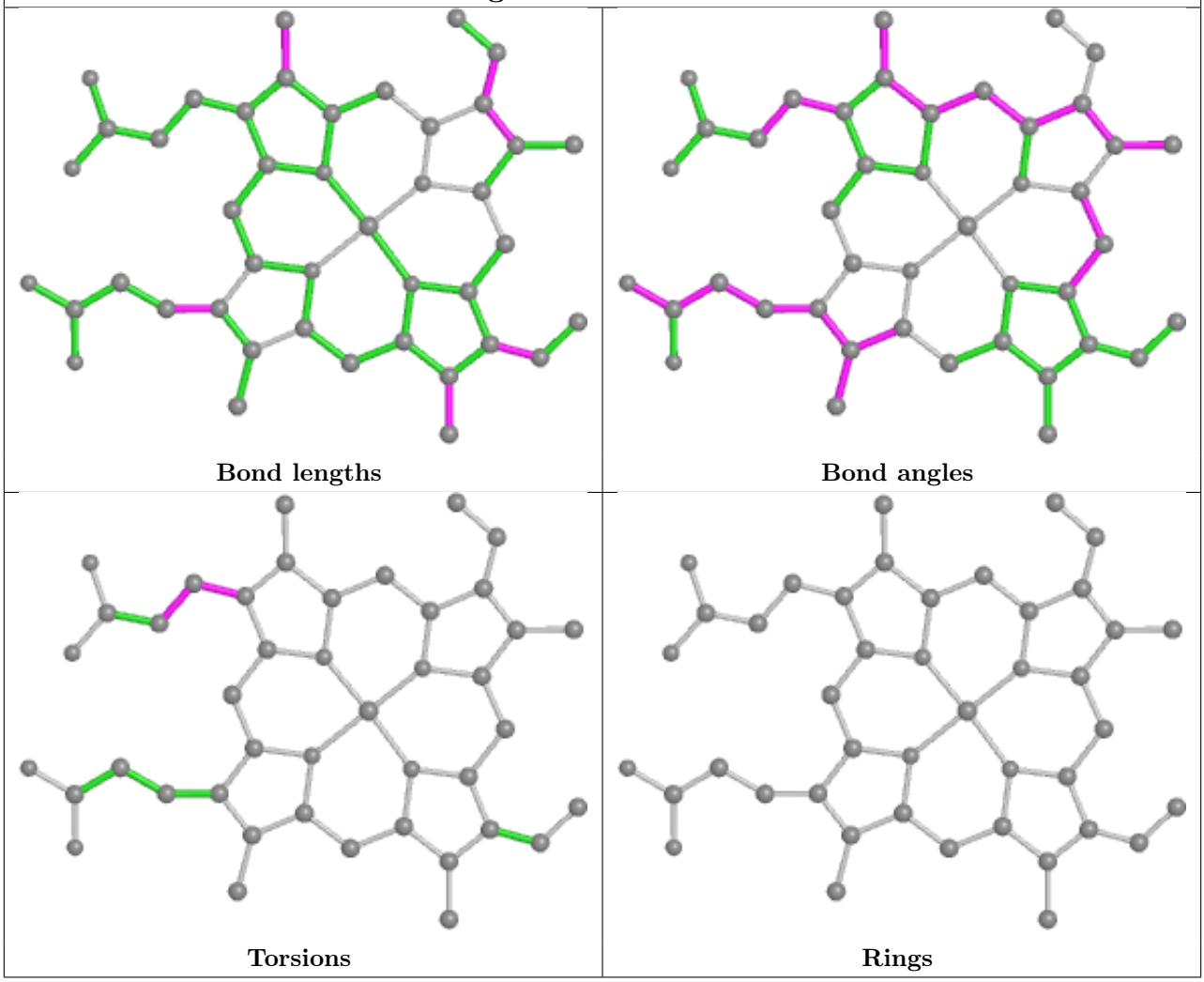


Torsions



Rings

Ligand HEM B 200



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	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	119:PRO	C	120:ALA	N	1.17

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.